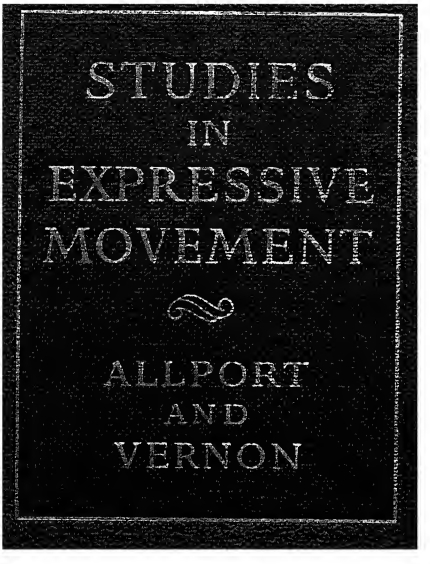

STUDIES
IN
EXPRESSIVE
MOVEMENT
ALLPORT
AND
VERNON

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN COMPANY
OF CANADA, LIMITED
TORONTO

STUDIES IN EXPRESSIVE MOVEMENT

BY

GORDON W. ALLPORT, PH.D.

ASSISTANT PROFESSOR OF PSYCHOLOGY,
HARVARD UNIVERSITY

AND

PHILIP E. VERNON, M.A., PH.D.

FELLOW OF ST. JOHN'S COLLEGE,
CAMBRIDGE UNIVERSITY

With a chapter on Matching Sketches of Personality with Script, by

EDWIN POWERS, A.M.

DARTMOUTH COLLEGE

New York
THE MACMILLAN COMPANY
1933

Published January, 1933.

SET UP, ELECTROTYPED, AND PRINTED BY T. MOREY & SON
IN THE UNITED STATES OF AMERICA

INTRODUCTION

Investigations of personality may be focused upon any one of three different levels of phenomena. The first is the level of traits, interests, attitudes, or sentiments considered as composing an "inner" personality; the second is the level of behavior and expression; the third is the level of impression, the perception and interpretation of behavior by another. Since a discovery on one of these levels establishes a presumption that the phenomenon in question has some counterpart on the other levels, a problem which is elusive on one plane may often be more expediently attacked on another. This is the motive and the plan behind the present study. Instead of approaching the difficult problem of consistency or organization in personality through a study of "inner" dispositions—which, of course, can only be known indirectly through tests and scales,—we have chosen to refer the problem to the level of expressive movement and there to examine it in a more direct fashion.

Besides being expedient, this policy of referring knotty problems of personality to the field of expression has considerable theoretical justification. Unless we accept the epistemology of intuitionism, we are forced to regard our judgments of personality as inferential constructs based upon our sensory perception of expression; and to assume that it is only through our perceptions of the physical bodies, speech, or gesture of our associates that we derive any knowledge of their natures. From this point of view the direct study of expression is the most natural possible approach to the study of personality.

This statement is not intended to imply that the problem of expressive movement is coextensive with the problem of personality. There are innumerable questions concerning the dispositions that lie behind movement and the effect of this movement upon observers. It would be misleading, for example, to claim that whatever consistency we may find in expression must have an exact counterpart in the latent dispositions of personality or in the perceptions or impressions of others. Köhler (*89* pp. 261 ff.) and Ichheiser (*76, 77*) have shown that the problem of the interrelation of these three levels is very intricate; although each in itself shows a certain structured quality, the central, the motor, and the phenomenal organizations may not strictly correspond. Ichheiser believes, for example, that there is much greater unity in the impressions gained from a person's acts than there is among the acts themselves, and he gives reasons why the "ethical" (dispositional) self and the "aesthetic" (expressed) self may likewise be at variance.

Hence, in approaching consistency from the side of expression, we cannot claim to cover in full the problem of organization of personality. Consistency may be, and actually has been, studied on both the other levels. Research on the internal consistency of attitude scales and tests for traits attack the problem from the point of view of "inner" dispositions. The experiments of Wolff and Arnheim (cf. pp. 11–15) study unity from the side of impression. A synthesis of all these approaches would contribute much to an understanding of the nature of organization in personality, but in the present preliminary state of research, investigators may be excused for preferring to confine themselves to one level. The failure of experimenters and writers, however, to recognize other levels than the one on which they work, often leads to

restricted and inadequate definitions of personality. We find, for example, definitions exclusively in terms of the inner self, others that take account of nothing but behavior, and still others that consider only the effect of conduct upon other people. The experiments reported in this monograph are based largely upon the investigation of a single level; however, in our discussion of the results (especially in Chapters VI and VII) we will consider to some extent the implication of our work in the wider problem of the organization of personality as a whole.

Since the purpose of the study is not primarily to determine the correspondence between peculiarities of expression and the inner dispositions of personality, it must not be looked upon as an investigation in "character analysis." At most, it is merely a critical propaedeutic to the elusive problems of psychodiagnosis. If our experiments should reveal a considerable degree of consistency among an individual's expressive acts, then the question of the ultimate validity of the arts of physiognomy and graphology may profitably be raised. But if there is no ascertainable consistency among the fields of expression, it would seem somewhat fruitless to worry about their interpretation in terms of inner traits.

Expressive movement in this monograph refers to those aspects of movement which are distinctive enough to differentiate one individual from another. Most movements have both non-expressive and expressive features. To illustrate, the eye blink reflex is a simple phenomenon of adaptive movement common to all men, and insofar as it is adaptive and common, it has no expressive significance; but as Ponder and Kennedy (*132*) have shown, there are individual *manners* of blinking. The analysis and comparison of these individual manners of blinking

would, therefore, constitute a problem in expressive movement. It is true that the term "expression" is usually employed in psychological writing in connection with emotional behavior, and also that it does not always refer to individual differences. There is, however, great need for a term that will stand for *individuality of expression*, not only in emotion but in all motor activity. With this explanation the use of the term in our special sense should give no difficulty.

The reader may feel that this definition of expressive movement leads us directly into the recognized field of the psychology of individual differences. It does; but it also leads us beyond, into the psychology of personality. These two fields are by no means identical. Dodge (*36* p. 6), commenting on the great advance of differential psychology over the older psychology of the "average" man, writes as follows: "Treating each individual as a special combination of capacities, accomplishments, and tendencies, has been far more productive than treating individuals as though they were all alike or as though they belonged to mutually exclusive types." In point of fact, however, the psychology of individual differences has *not* treated the individual as a *special combination* of capacities, accomplishments, and tendencies; indeed it has not, strictly speaking, treated differences between *individuals* at all, but merely differences between reaction times, fatiguability, intelligence, or similar functions abstracted from the total mental life of the individual. The utmost that differential psychology has accomplished toward the study of personality is the construction of graphs which simply list or plot scores for these separate functions. The special combination or patterning of these attributes and their coherence and consistency with one another have seldom been considered. The problem of

the *organization of individual functions* has always eluded differential psychology.

The present study, then, goes beyond the tradition of the psychology of individual differences; it treats the question of the agreement among individual differences, with special reference to the motor field. The discovery of well-integrated and consistent expression in the motor field would establish a presumption that similar patterning is to be expected in all aspects of personality, but it is for the future to show the exact relation of expressive movement to "inner" traits and to impression.

G. W. A.
P. E. V.

December, 1932.

ACKNOWLEDGMENTS

With the exception of the work on handwriting reported in Chapter X, the material for this volume was gathered chiefly from experiments conducted in 1930–31 at the Harvard Psychological Clinic. Dr. H. A. Murray, director of the clinic, provided facilities for the investigation and gave valuable advice. Our co-workers were Mr. Otis Rice and Mr. J. H. Smith. In addition to the professional graphologists whose coöperation Professor Powers acknowledges in Chapter X, the following persons gave both time and skill to judgments of handwriting: Professor J. E. Downey, Dr. R. E. Saudek, and Mr. DeW. B. Lucas. Assistance in various other portions of the work was given by Professor E. G. Boring, Dr. H. Cantril, Mr. H. E. Odbert, Mr. H. A. Rand, Mr. R. N. Sanford, Mr. C. E. Smith, and Dr. J. Volkmann. For financial assistance in making the results available for publication the writers are indebted to the Pierce Fund of Harvard University and to the Social Science Research Council. Finally, there were the numerous subjects whose interest and faithfulness made the experiments possible. To all these generous collaborators the authors feel deeply grateful.

Through the death of Professor June E. Downey, which occurred while this monograph was in press, psychology, especially that branch which deals with the study of expression, has suffered a great loss. The reader who is familiar with her important researches in the psychology of handwriting will recognize that the second part of this volume rests largely upon foundations which she established. It was characteristic of her work that while keeping within the standards of scientific experimentation, she managed with unusual success to devise and to apply methods which did not distort the problem under investigation. She has shown that the psychologist who studies personality may make richer discoveries if he is aided in his scientific approach by his aesthetic sense and by his naïve understanding and appreciation of human beings.

TABLE OF CONTENTS

CHAPTER I

THE PROBLEM OF CONSISTENCY IN EXPRESSIVE MOVEMENT

> Every thing in man is progressive; every thing congenial; form, stature, complexion, hair, skin, veins, nerves, bones, voice, walk, manner, style, passion, love, hatred. One and the same spirit is manifest in all.
>
> — LAVATER

Literary and Speculative Approaches. The most famous of physiognomists, Johann Kaspar Lavater, considered the basic principle of expression to lie in the fact that no feature of the body contradicts any other feature. It is not, he assures us, the eyes that laugh nor the shoulders that laugh; it is the *man* that laughs. Lavater was not an experimentalist, and according to modern standards did not offer proof of this fundamental theorem of physiognomy. Even though common sense may dispose us favorably toward Lavater's position, it still remains in the absence of accurate and controlled study, sheer dogma. And if like many dogmas this one contains both truth and falsehood, nothing excepting analysis and measurement can tell us the proportion of each.

Essayists, more frequently than experimentalists, have reflected upon the consistency of the expressive acts of an individual, and their significance in representing his personality. C. H. Cooley, for example, is impressed by the interdependence of an author's modes of expression. He writes

> Can anyone . . . read Macaulay and think of a soft and delicately inflected voice? I imagine not: these periods must be connected with a sonorous and somewhat mechanical utterance; the sort of person

that speaks softly and with delicate inflections would have written otherwise. On the other hand, in reading Robert Louis Stevenson, it is impossible, I should say, not to get the impression of a sensitive and flexible speech. Such impressions are mostly vague and may be incorrect, but for sympathetic readers they exist and constitute a real, though subtle, physiognomy (*30* pp. 112 f.).

In spite of Cooley's warning that "such impressions are mostly vague and may be incorrect," the belief that the style of expression is the signature of personality, always consistent with itself, is very prevalent. In an essay on Eugene O'Neill, Barrett Clark describes certain "extraordinary" physical features and expressive gestures of the playwright in order to make more intelligible the "extraordinary" quality of his work. In his preface to the letters of Franz Schubert,[1] Ernest Newman writes, "Schubert's letters are a true counterpart of his music: the style is simple—melodic and diatonic, we may almost call it—without any involutions or complexities . . ." Newman tells us that "the uncouth insufficiency of some of Beethoven's first sketches for a musical work has its counterpart in the general shapelessness of his literary style," and that "the flamboyant exaggerations of the earlier Berlioz are reproduced in his prose." He points also to the obvious parallel between Wagner's musical compositions and his "highly involved prose, with its maddeningly long sentences, its syntactical involutions, its enclosure of one parenthetic clause in another . . ."

In our program notes at the symphony, we read concerning Handel that "his personality and his music alike were characterized by the same ineffable majesty." Biographers and critics seem to consider it only natural that the diverse expressions of a single personality should show striking similarity.

[1] *Franz Schubert's Letters and Other Writings*, edited by Otto F. Deutsch, London, Faber and Gwyer, 1928.

Psychologists themselves have occasionally been interested in this problem, though seldom has the interest led to direct experimentation. Implicitly, Woodworth's "adverbial" definition of personality (*202* pp. 552 ff.) contains a statement of the issue. If, as he says, personality is revealed in the *way* in which an individual talks, remembers, thinks, or loves, the question naturally arises as to the degree of uniformity in the individual's manner of doing all these things. More direct is Roback's (*142*) discussion of the "personal idiom," wherein he, like Newman, shows how the same quality often pervades the different creative works of an artist. Roback (*141*) and Krout (*93*), like Freud (*55*), are interested in the symbolic value of gestures. Idiosyncrasies of movement, they hold, proceed from complex inner dispositions and can never be treated as isolated motor phenomena.

In a study of the voices of nine speakers over the radio Pear (*130*) approaches our problem very closely. Many of his listeners ventured to give physical descriptions of the speakers who were known to them only from their reading over the radio of a passage from Dickens. Although these descriptions were difficult to score, they seemed on the whole to be more right than wrong. In one case, where the speaker was a detective-sergeant he "was almost invariably described as being a robust man of heavy build, stout and burly." In another connection Pear observes that, "since speech is a form of action, it is natural that some persons' speech-style should reflect their *general* behavior." Again, "in our estimation no speaker's appearance was at variance with his or her voice." These last two observations, though unsupported by experimental evidence, are at least interesting statements of our problem. Suggestive illustrations of consistency are found also in the writings of the criminologist,

Hans Gross (*62*), and an eminently practical test of the matter is promised in the *modus operandi* system of classifying crimes (*188*). According to this system it is assumed that each criminal employs methods peculiar to himself which if analyzed and recorded will aid in identifying him as the malefactor in a new case.

None of these antecedent studies gives our problem an altogether satisfactory formulation. Most of them are subject to criticism for their confusion of two quite distinct issues. In the first place, there is the problem of the meaning or *diagnostic significance* of expression (*i.e.*, whether an idiosyncrasy in manner is a valid indicator of some personal complex, prejudice, or interest). Secondly, there is the problem of the *inter-agreement* of an individual's various expressive habits (*e.g.*, whether the heaviness of Macaulay's prose actually did have some recognizable counterpart in his vocal expression). It is clear that the first problem cannot be satisfactorily attacked until some solution of the second problem is available. Unless we can discover what statisticians call the *reliability* of our indicators, we should not venture by the aid of these indicators to diagnose inner qualities of personality. For example, before attempting to verify Friedrich Gerstäcker's character-readings from a man's manner of wearing a hat,[1] it is necessary to know whether the position of the hat is a characteristic one for the wearer. It is necessary likewise to know whether the man who wears his hat, say, on the back of his head exhibits other movements supposed to be indicative of laxness and extravagance.

[1] "If he wears it perpendicular, he is honest, pedantic, and boresome. If he wears it tipped slightly, he belongs to the best and most interesting people, is nimble-witted and pleasant. A deeply tipped hat indicates frivolity and an obstinate imperious nature. A hat worn on the back of the head signifies improvidence, easiness, conceit, sensuality and extravagance; the farther back the more dangerous is the position of the wearer. The man who presses his hat against his temples complains, is melancholy, and in a bad way."

Experimental Approaches. Many American investigators have crusaded against the popular type of physiognomy which claims to establish parallels between single physical features and traits of personality. However serviceable to the public as a warning their iconoclasm may be, it has led to little constructive research, and has contributed nothing to a psychological understanding of the tangled issues involved. On the contrary, it has actually obscured legitimate problems within the field of expression. It is not uncommon, for example, for American psychologists to classify graphology with astrology, and to condemn both with one breath. In so doing these psychologists simply confuse a *bona fide* problem of motor psychology with occultism. Again, the passion for correlating isolated physical characteristics with some ill-conceived "trait" blinds many investigators to more promising avenues of research.

One of the most widely cited investigations is that of Cleeton and Knight (*28*). Although useful as a caution against popular physiognomists, it was not designed to advance the serious study of expression. For example, the fact that seventy judges made reliable judgments based on a few minutes' observation of the behavior and appearance of the subjects is too lightly passed over. The writers were interested in showing that these judges could not have based their decisions on the legendary signs employed by certain popular physiognomists; but the challenge of the fact that some other aspects of expression are interpreted consistently by judges is not recognized.[1]

[1] For our purposes, of course, it is irrelevant that the ratings by these judges did not agree closely with ratings made by acquaintances of the subjects. Our point is that they agreed well enough among themselves to show that consistent impressions are gained from a few moments' observation of physique and activity.

Contrasting with the majority of American studies, research on the Continent takes for granted that no case can be made for the older physiognomy, and proceeds hopefully to new and rather daring formulations. From the Institute of Child Research at Leningrad, for example, Oseretzky (*127*) has issued a comprehensive plan for the study of *Psychomotorik*. In his scheme this large field is subdivided into three sections. *Motoskopie* aims to analyze and classify all the significant types of expressive movement (posture, pose, facial expression, gesticulation, handshake, gait, speech, handwriting, automatic movement, pathological movement). *Motometrie* is concerned with the measurements of movements, and *Motographie* with their recording. Actually the chief aim of this Russian work is the development of an age scale for the determination of the motor maturity of children. Unfortunately for our purpose, the interagreement of the test items in this scale is not given, so that no light is shed upon our problem of motor consistency. We are, however, indebted to Oseretzky for the broad perspective he gives to the whole subject of expression; several suggestions from his *Motoskopie* are included in our survey of the forms of expressive movement on pp. 24–35.

The work of Enke (*48*) on psychomotor types is directly concerned with the problem of consistency. His purpose is to determine whether there are features of movement which differentiate Kretschmer's pyknic and non-pyknic types. His findings are positive: (1) whether the times are measured by spontaneous tapping or by the congenial pace of work, pyknics are essentially slower in their movements than leptosomes and athletics; (2) in tests for freedom of movement pyknics are more irregular and variable, the athletics and leptosomes are more me-

chanical, automatized, and stereotyped; (3) leptosomes and athletics are given to perseveration, *i.e.*, they find it difficult to change their personal rhythm, and are slowed by distraction; (4) pyknics are adaptable to external rhythms and their action is speeded by distraction; (5) finely coördinated activity is best carried out by leptosomes, less well by pyknics, and least well by athletics; (6) pyknics are "fluid, free, soft, rounded, uninhibited" in their actions, leptosomes are "stiff and angular"; (7) pyknics fatigue gradually, leptosomes suddenly; (8) pyknics have a lower average handwriting pressure but greater variability than do non-pyknics. Taking the results together, Enke concludes that the movements of pyknics tend to be slow, free, adaptable, uninhibited, easy-going but variable. The schizothymic temperament which is associated with non-pyknic physiques seems, on the contrary, to express itself in movement that is hesitant, cautious, critical, tense, stereotyped. For our purposes, of course, the important point is that these results were arrived at on the basis of the inter-agreement of several tests. According to Enke it does not matter, for example, whether the typical leptosome is writing with a pen, carrying a glass of water, or reacting to music, he is found to be uniformly tense and cautious.

The present investigation, like Enke's, administers a large number of motor tests and examines their agreement with one another. In other respects, however, the methods of these two studies differ. Enke employed 500 subjects; the present study, 25. But Enke's advantage in numbers is somewhat offset by his failure to determine the reliability of his measures. From his results it is possible to conclude only that, by and large, pyknics and leptosomes differ in various aspects of expressive movement; the amount of the differences is stated merely in

percentages (without reference to probable errors). Enke entirely omits the problem of personal (intra-individual) consistency. From his data it is impossible to tell whether a *given* leptosome possesses all, some, or even none of the expressive habits discovered to be characteristic of his type. The present study, on the other hand, makes no use whatever of typological assumptions. It is interested only in the question whether each of our 25 single subjects is consistent with himself.

Another German publication possessing sweep and originality is *Gang und Charakter* (*18*), a collection of essays submitted in a prize competition. Only one of the contributions (Wolff's) is genuinely experimental, and must be considered separately. The rest are for the most part programmatic and critical. In Wilsmann's essay, for example, we find gait analyzed into one subjective or qualitative characteristic (rhythm) and seven objective or quantitative characteristics (regularity, speed, weight, length, elasticity, definiteness of direction, changeableness). Each of these objective features, he claims, needs to be interpreted in the light of the unanalyzable factor of rhythm, a procedure very similar to Klages' in the field of handwriting.[1] Most of the contributors to *Gang und Charakter* seem to hold with Wilsmann that there are unanalyzable features in gait, and that experiment is foredoomed to failure.

Wolff's study (*18*), on the other hand, is close in spirit

[1] *Rhythm* is a term frequently encountered but seldom defined in literature on expression. It seems usually to stand for an "indefinable something" which provides a refuge for those who are unwilling to surrender movement entirely to analytical and quantitative treatment. In this sense "rhythm" is something that can only be *understood* by "intuition"; it is a pattern too intimate and integrated and too meaningful to be studied adequately through analysis. Chapters X and XI of the present study deal with the "rhythm" or total pattern of handwriting as an expression of the unanalyzed personality; but in Part A, the concept is employed as little as possible. When it appears it usually means merely periodicity; if any other meaning is intended the context makes it clear.

and in conception to our own. His statement of the methodological dilemma facing investigators is worth quoting:

> This is the standpoint of characterology to-day: Either we have followers of the statistical-mechanical procedure which yields a crude succession of single elements which never give a living picture; or else we have the intuitive penetration which, it is true, can delineate brilliantly the living processes of human beings, but cannot be communicated, verified, or utilized for a knowledge of genetic relationships (*18* p. 109).

To resolve this conflict between a scientific but inadequate characterology and an intuitive but incommunicable characterology, Wolff proposes the "experimental-dynamic" methodology which involves

> (1) a surrender of the "total" personality, but a retention of interest in complex and structured forms of expression (to be studied either by themselves or in relation to one another);
>
> (2) a standardization of outer situations, so that the resulting differences of behavior may be referred only to variables in personality;
>
> (3) an exclusion of subjective factors such as those resulting when the experimenters have unequal acquaintance with the subjects;
>
> (4) the use of several judges, when judges are necessary, to control reliability;
>
> (5) the study of movement that is spontaneous and unaffected by self-consciousness of the subject.

In an early experiment (*200*) Wolff obtained five sets of data from each of his subjects: (a) the subject's vocal expression recorded without his knowledge, (b) a profile picture of the subject's face, (c) a specimen of his handwriting, (d) a picture of his hands, (e) his manner of retelling a folk-story, stenographically recorded. Wolff endeavored to determine whether judges could match correctly the various records taken from any single subject. His principal results may be summarized as follows: (a) profiles and hands were not matched more frequently than chance would allow with handwriting, voice,

or *Nacherzählung;* (b) voice records and handwriting were matched correctly about one and a half times as frequently as would be expected by chance; (c) the same frequency obtained in the matching of specimens of handwriting with styles of retelling stories; (d) when a judge was not required to match records but to characterize them, his descriptions of the different records from a given subject tended to be identical, (for example, if the retelling of a story was described as "colorless," the handwriting of the subject was also likely to be described as "colorless" by the same judge). Wolff is disinclined to give a more strictly statistical statement of his results, for statistics, he holds, cannot express adequately all the consistency discovered by his method.

In his experiments on gait he employed the moving picture camera. His subjects were mostly students, of both sexes, dressed in a standard, loose fitting garment. He assigned them the simple task of picking up rings and tossing them over a stake. A week after the films were made they were shown to eight judges who knew the subjects and to five who did not. The heads in the films were blocked out, and the judges who knew the subjects were first asked to identify them if possible. Of these indentifications 30% were correct and 70% false. The subjects, who also acted as judges, were able in 100% of the cases to make correct self-identifications. This last result contrasts with Wolff's previous discovery that, from records of voice, hands, profile, and retelling, identification of others was easier than identification of self. From this curious result it appears that one's expression is recognizable to oneself only when it involves the large muscles of the trunk and limbs.

The judges were also asked to describe each personality from gait. No validation for these descriptions is given,

but Wolff notes that many of them correspond exactly to characterizations given for the same subjects in his earlier experiments. That is to say, the impression made by the gait of a given subject was likely to be reported in the same terms as the impressions made by his style of narration, handwriting, or voice. This fact is perhaps more remarkable since the experiments were separated by an interval of two years. No quantitative statement of this identity of descriptive terms was attempted. Wolff, however, has made a definite experimental attack upon the problem of intra-individual consistency in expressive movement. With the aid not of impression but of correlation we hope in the present study to improve upon Wolff's method, if not upon his clear-cut perception of the problem.

A study employing a "matching method" very similar to Wolff's is that of Arnheim (7). Starting with a simple discovery by Wertheimer that the handwritings of Michelangelo, da Vinci, and Raphael could be rather accurately identified by observers who had no acquaintance with the artists' chirography, Arnheim proceeded to invent numerous problems of a similar order. Employing over 75 sets of material, and 65 judges, he assembled voluminous data on the consistency of expression. Quotations were matched with photographs, handwritings with sketches of personalities, silhouettes with single descriptive terms, and so on. The matchings were made in all variety of combinations (*e.g.*, three photographs with one, two, three, four, or six sketches). It is difficult to present Arnheim's very miscellaneous results in any standard fashion. In practically every case, however, the number of correct matchings exceeds chance, often by a very large margin. In only two experiments was the number right less than probability, and in only one case was the

number wrong greater than chance. Dividing his "half right" answers equally between the rights and wrongs, we may summarize rather roughly some of his findings as follows: In all experiments in which there were two choices, the average number of correct matchings was 70.9% instead of the 50% expected by chance; where three alternatives existed the results were more favorable (83.6% instead of 33.3%); where six alternatives existed for a given standard, correct matchings were 51% (instead of 16.7%). Arnheim claims that for all the experiments together the average correct matching is about two and a half times that expected by chance.

It should be noted that although both Arnheim's and Wolff's experiments are explicitly aimed at intra-individual consistency, their matching methods depend largely upon the factor of *subjective impression.* Having discovered consistency in *impression* they conclude by inference that there is a corresponding consistency in *expression.* Excepting for the graphological studies (Chapters X and XI), the present investigation employs almost entirely *direct objective measurement of consistency.*

It seems that the only American investigation which closely resembles our own is that of Downey (*39*). In 1919 she published an account of a study designed to test the proposition that "graphic individuality is but a specific example of a pattern that is impressed upon all the expressive movements of a given person" (*39* p. 97). She herself classified the handwriting of 12 subjects according to her judgment of the scripts as rapid or slow, light or heavy, loose or compact, expansive or restrained, adroit or maladroit, fluent or jerky, angular or supple, conventional or individual, impulsive or deliberate, concentric or eccentric. Eleven judges were asked to classify the same 12 subjects under the same rubrics, basing their

judgments not upon handwriting but upon their impression of the subjects' gait, carriage, and gestures. The method of scoring was simple. If the majority of judges chose the same rubric for a given subject as Downey did, an "agreement" was recorded. Chance would allow 50% agreement; the actual correspondence is 60.5%. Her final conclusion is cautious: "It may be stated that the outcome of the experiment is slightly in favor of an agreement between graphic and expressive movement. . . . Certainly no sweeping assertions of general similarity can be ventured, although for a few traits there is strong evidence of such harmony" (*39* p. 104). If the three pairs of rubrics which gave the greatest subjective difficulty, viz., light—heavy, individual—conventional, adroit—maladroit, are eliminated, her results indicate more than 70% agreement.

Downey's method is more controlled than Wolff's and Arnheim's, since her conditions excluded the free "intuitive" perception of similarities in different fields of expression. In her work there was no direct matching of two sets of records by a single judge, but rather two entirely independent sets of judgments which were compared objectively. Likewise she allowed no freedom for spontaneous characterization of the record; all her terms were prescribed in advance, and insofar as they were ambiguous or not understood they affected her results adversely. Finally, she did not use complete records of expression but only ten "traits" of movement, so that we cannot say from her results whether or not handwriting *as a whole* shows similarities with carriage and gesture considered *as a whole.* It is quite possible that all these conditions, though making for control, resulted in a method that was after all unsuited to the problem she was attacking.

In three respects Downey's experiment anticipates our own. (1) She formulates precisely the same problem, though her method gives relatively more prominence to graphic movement. (2) She has evolved an objective method for what at first sight seems to be a hopelessly subjective problem. Her technique resembles somewhat that which is employed in Chapter XI of this monograph. (3) Her examination of individual records shows that even when her judgments and those of her collaborators seem hopelessly at variance in a quantitative sense, there are frequently meaningful connections between them which elude statistical statement. This last point anticipates our own discovery of "congruence" which is discussed in detail in Chapter VI.

The Genetic Approach. Popular lore has helped us to identify our problem, and a few antecedent experiments have provided us with suggestions of method. But the issues with which we are dealing appear more concrete and psychologically more arresting when we regard them from the genetic point of view. The developmental approach shows clearly that there is a *theoretical* expectation of consistency in adult expression.

In early infancy skeletal movement is crude but strikingly unified. Lewin (*100, 101*) has demonstrated that in babies of six months mass-action of the gross musculature predominates over specific reflex action. It is the entire body, and not the single limb, which reaches, retracts, or expresses fretfulness, expectancy, or contentment. This early or "protopractic" consistency has been reported by many other observers. Allen (*1*) finds that some infants as old as one year reach with both hands (that is, express the same tension with two limbs) when one would suffice. Halverson's cinematographic studies of prehension (*65*) show the same type of phenomenon.

Pratt, Nelson, and Sun (*134*) as well as Coghill (*29*) offer proof that the course of development in general is from the diffuse responses of the whole body to the more specialized expression of specific organs or "segments."

Obviously, however, when the child acquires greater precision in the manipulation of limbs and voice, this diffuse manner of adjusting recedes. (There are, of course, occasional regressions during strong emotional excitement.) Differentiation sets in; limbs, voice, and facial muscles no longer necessarily act as a unit. Mere knitting of the brows in the adult may signify as much pain as random thrashing and screaming in a child; a slight shake of the head may express as definite an avoidance as the retraction of the infant's whole body.

However, before assuming that this growing differentiation leads to complete specificity in the expressive movement of the adult, it is necessary to consider several counter-influences. For one thing, the original gross movements of the infant often have a distinctive *general* character that seems to persist even while the specialization is in progress. Consider, for example, the case of Infant *A* and Infant *B*, both four months of age. *A* is excitable and aggressive in his crude and random actions; *B* is sluggish and passive. Four years later these children are seen together again. Although both have grown immeasurably in skill and precision in the use of the limbs, their specific acts still show the same "temperamental" differences. This is a common observation, although it seems not yet to have been experimented on successfully. From the Murphys' survey of the problem of consistency of traits with advancing age (*118*, pp. 210 ff.; 243 ff.) it appears that attempts up to the present time have employed too few children, have covered too short a span of time, or have used scales too unreliable to yield

conclusive results. But Bühler (*24*) and Zoepffel (*206*) have definitely established the existence of such "temperamental" differences in the behavior of children only a few weeks old; some are nervous, others placid and dull, some socially responsive, others socially unresponsive. Moreover Washburn (*192*) finds that babies from two to twelve months are markedly consistent in respect to the amount and manner of their laughing and crying. Therefore, it appears that we must allow for the influence of certain general psychomotor factors which operate to maintain a "temperamental" consistency in behavior during the course of increased specialization in skill. To borrow Dodge's terminology (*36* pp. 96–100), we must recognize that certain qualities of the original "protopractic" mass-action seem to remain in each of the new "epicritic" responses.

Likewise, we must remember that differentiation is never complete. Emotional states even in the adult are "patterned" in their expression, and as Landis has shown these patterns are consistent in any individual case. Kempf (*83*) believes that not only emotion but also postural tonus is diffused and relatively undifferentiated. A man's posture, he argues, should be regarded as a total organic response which represents an equilibrium between the autonomic tensions and the demands of the environment. The "arrogant" postures of La Bruyère's *Giton*, who not only sat down heavily but blew his nose loudly and spoke like an oracle, or the self-effacing manner of *Phaedon*, who not only blushed when he spoke but even spat unsuccessfully, are illustrations of such a general psychomotor tonus.[1]

In yet another sense specialization in expression fails

[1] The sketches by the famous "Character" writers frequently depend for their effect upon descriptions of psychomotor consistency. Cf. Richard Aldington, *A Book of Characters*, N. Y., E. P. Dutton.

to become complete. Much of the behavior in the young child is synkinetic; that is to say, his adaptive acts are accompanied by auxiliary movements. This seems to be a natural stage in the process of transition from adjustment with the whole body to adjustment with single limbs. In fact, Oseretzky (*127*) finds these *Mitbewegungen* a serviceable index of inferior motor development. Sometimes, as Holt has shown (*71* pp. 225 f.), vestigial infantile movements persist throughout life and become motor idiosyncrasies of the individual. Many of these idiosyncrasies, of course, are quite specific and persist merely as "symbolic gestures"; but they are at the same time further evidence that adult expression is not completely specialized or entirely free from its early ties.

Finally, in the course of development there are personal and professional influences which tend to offset the process of differentiation. Integration and conventionalization forge new unities. A child, for example, leaving the stage of "protopractic" consistency, learns to use his hand and tongue with precision and independence. In manual training he acquires an undistinguished formalized style of work, and in his English course, a standardized style of composition. But with time these immature and unrelated expressive habits seem to acquire a common character. *Both* his handicraft *and* his diction lose their school-like qualities, and seem to acquire a new similarity; they may both be restrained, or orderly, or ostentatious. Habits that have developed independently appear finally, when the personality is standardized, to become consistent with one another. It is Ruttman's belief (*146*) that with advancing age there is greater uniformity between expressions of the tongue and the hand. These two members are, *par excellence*, the agents for the expression of personality, and as personality becomes more integrated

and unified, the manual and lingual expressions cannot help but become more consistent with one another. The pedant seems to become more and more pedantic in speech and gesture; the soldier, more and more soldierly.

The genetic approach, then, calls attention to two facts: in the beginning of life, movement shows a striking degree of unity; this unity is not completely broken down by the process of specialization and differentiation. Accordingly, from the genetic point of view, there is a strong presupposition in favor of our discovering consistency in the expressive acts of adults. We might add that this presumption is strengthened by a number of studies, reviewed by Schwangart (*161*), which show that consistent motor individuality is found among animals.

At the same time, since differentiation and specialization are inevitable processes tending to offset the original unity and the progressive integration of movement, the investigator must be prepared to find the problem of consistency far more complex and intricate than the problem of mass-movement in infants. Everyone knows, for example, that the adult's body is capable of considerable self-contradiction. The eyes may betray boredom while the voice expresses interest; a frigid handshake may belie a friendly smile; a man's gait may seem self-assured while his speech is hesitant and embarrassed. Many of the contradictions in expression vanish, however, when one perceives them as reflections of a single, complex state of mind. It may be that the man in our illustration *is* both friendly and unfriendly; that the other *is* both self-assured and embarrassed.

This situation leads to a fundamental distinction of meanings in the term "consistency." There are clearly two senses in which movement may be said to be "consistent." On the one hand, the term suggests that dif-

ferent indicators of movement must vary directly with one another; on the other hand, it suggests that even when indicators do not agree directly with one another, they may still be harmonious in the sense that they both express different aspects of a single complex state. In either case there is justification for the term "consistency," and, accordingly, in this study the term will be used to refer to these two phenomena collectively. For precision, however, we must distinguish between them. When in our results indicators are shown by correlation to vary directly with one another, they will be said to *correspond;* indicators that do not correlate, and yet are clearly related to one another through the complex medium of personality, will be said to be *congruent.*

A Tentative Analysis and Partial Classification of Expressive Movements. Several difficulties arise in attempting to define and classify expressive movements. One rather obvious difficulty lies in the fact that expression is not always confined to activity. Postures or cast of countenance are not, strictly speaking, movement, and yet they afford some of the best material for the study of expression. Another problem arises from the fact that the peripheral aspects of movement and posture cannot be distinguished sharply from mental attitudes, traits, or interests. And furthermore it is difficult to separate the expression from the resulting impression,—the actual from the interpreted expression.

To add to our dilemma no single act can be designated exclusively as "expressive," and none exclusively as "non-expressive." Every act seems to have its non-expressive as well as its expressive aspects. It has its adaptive or *zweckmässig* character, and also its individual or *ausdrücklich* character. In unlocking a door, for example, the task itself prescribes definite coördinated movements

suited to the goal, but it allows also a certain play for individual style in executing the prescribed movements. There are peculiarities in the steadiness, pressure, precision, or patience with which the task is executed. It is only these individual peculiarities that are properly called "expressive."

Finally, even these peculiarities cannot all be regarded as expressive of *personality*. Many of them are influenced by external conditions which are quite unrelated to personality. The following factors, for example, over and above the integrated systems of habits and traits which we call personality, are important determining influences:

(a) the exigencies of the immediate goal
(b) pathological or accidental deformation of the body
(c) conditions of health and disease
(d) individual peculiarities of muscular structure or bodily metabolism
(e) constitutional make-up (cf. Enke's *48*, Yizlin's *203*, and Gurevich's *63*, work on differences in motility among Kretschmer's biophysical types)
(f) age
(g) sex
(h) strain and fatigue
(i) conditions of the physical environment (*e.g.*, ground and climatic factors in walking; pen, ink, and paper in handwriting; the clothes, and shape of chair in posture)
(j) transitory emotional states or moods
(k) racial tradition
(l) convention or fashion
(m) special habits springing from special training (*e.g.*, elocution, dramatics, athletics)
(n) temporary social environment, leading to artificial manners or to a masking of normal expression.

Since several of these factors probably play a part in the determination of every act, we cannot assume that consistency in movement derives exclusively from some corresponding consistency in personality. Certain of these influences (*e.g.*, disease) may operate to produce a spuriously high degree of consistency among our measures.[1] On the other hand, some of the factors, since they vary from situation to situation, would operate to produce a spuriously low correlation among our measures, and lead us to underestimate the true consistency among expressive acts. In an experimental investigation it is possible to eliminate many of these conditions, and wherever possible, in the present study, this of course has been done.

The practical effect of this recital of difficulties is to make us wary of defining too narrowly the field of expressive movement or of classifying too rigidly its phenomena. Perhaps all we can say by way of definition is that our study of expressive movement is concerned with *individual differences in the manner of performing adaptive acts, considered as dependent less upon external and temporary conditions than upon enduring qualities of personality.*

The chief purpose of the schedule which follows is to call attention to the richness of an unexplored territory, and to provide a scaffolding for our own investigation. It is neither exhaustive in its scope nor minute in its analysis. Occasionally references will be given to works which offer more detailed information on the items listed, but nothing would be gained by attempting here to subdivide the field more finely. We would soon lose sight of

[1] The force of this criticism is somewhat weakened, however, by Giese's (*18*) observation that external conditions which constantly influence movement are likely also to influence and to form personality, and so, in the long run, to bring movement and personality into harmony.

our problem and be plunged into the absurd task of reducing expression to the contraction of single muscles.

The schedule attempts, without complete success, to keep separate those aspects of movement which are objectively measurable from those that can be described only in qualitative terms. In Part I A are listed a number of these qualitative terms which are frequently employed to designate *patterns* of movement as perceived and interpreted by others. The rest of the schedule contains chiefly, but not exclusively, terms referring to aspects of movement that may be objectively measured and compared. Our experiments are based upon a selection of fairly representative items from various portions of the whole list.

A TENTATIVE CLASSIFICATION OF EXPRESSIVE MOVEMENT [1]

I. General Features of Expressive Movement Which May Be Manifest during Any Activity

A. Descriptive Terms, Chiefly Qualitative, Commonly Applied to *Patterns* of Expression

1. Applied either to the total impression or to specific features

adroit, clever	maladroit, inept
artistic	crude
athletic	languid, effete
centrifugal	centripetal
ceremonial, formal, aristocratic	casual, informal, off-hand
confident, composed	nervous, tremulous
consistent, congruent	inconsistent, contradictory
conspicuous	colorless
coördinated	uncoördinated
dainty	clumsy
defiant	cringing, fawning
eager	abstracted, apathetic

[1] General References: Anderson (*6*), Downey (*40*), Gross (*62*), Krout (*93*), Lipmann and Bogen (*18*), Oseretzky (*127*), Perrin (*131*), Seelig (*163*), Vance (*185*), Ziehen (*205*).

eccentric, peculiar	normal, wholesome
energetic, vigorous	feeble
exact, accurate	careless
expansive	restrained, reclusive
explosive	stolid
extravagant, grandiose	unassuming
fluent, smooth	jerky, variable, erratic
forceful, commanding	ineffectual
frank	reserved
free, elastic, flexible	inhibited, rigid, inflexible
graceful	awkward, uncouth
impulsive, impetuous	cautious, deliberate
individual, unusual	stereotyped, mechanical
intense, decisive	hesitant, wavering
light	heavy
loose	compact
majestic, stately	simple
mature	childish
military	listless
naïve, natural	studied, stilted, affected, artificial
neat, precise	slovenly, sloppy, vague
purposive	aimless
quick	slow
refined	coarse, common
restless	calm, quiet, placid
rounded	angular
self-conscious	unself-conscious
strong, firm, emphatic	weak, indecisive
vivacious, lively, alert	sleepy, indifferent, inert

2. Applied chiefly to facial expression (or eyes)

absent	direct
ascetic	sensual, gross
benevolent	misanthropic, sardonic
expressive	expressionless
intellectual	stupid
penetrating	shifting
sad	cheerful
spiritual	worldly
sympathetic, kindly	hard, cruel, sinister
trusting	suspicious

B. Tonus [1]

constancy	irregularity
tenseness	relaxation
vitality	tonelessness

rhythm of innervation and denervation

C. Characteristic Poses

items in I A 1; also

sameness	variety
swift change	gradual change

1. Trunk and shoulders

chest thrown out, shoulders back
shoulders consciously squared
shoulders drooping
easy or palpitant respiration
regular or variable respiration

2. Head

bent forward	thrown back
drawn in	upright
inclined to side	mobile or stiff

3. Face [2]

items in I A 2, I E 4
characteristic wrinkles
flushed or pale
quickness in play of expressions
mask ability
regularity or asymmetry of contractions
predominance of one feature in facial play

4. Eyes

items in I A 2

avoiding	squinting
downcast	staring
shifting	straight forward
sagging lids	firm lids

[1] Theoretical aspects of tonus are considered by Kempf (*83*); techniques for its measurement are described in Chapter III.
[2] Cf. Dunlap (*46*).

5. Mouth
   items in I A 2
   asymmetric cast — pouting
   drooping corners — tight lipped
   twitching
6. Limbs
   items in I A 1, I E, II, III

D. Individual Differences in the Performance of Tasks
   items in I A 1
   accuracy — perseveration
   regularity — tempo
   orderliness — variability
   perseverance

E. Synkinetic or Auxiliary Movements (*Mitbewegungen*)[1]
   1. Few or many unessential movements while working or when idle
   2. Nervous and symbolic movements of hands
      fussing with objects, buttons, beads, ornaments
      twisting rings on finger
      pulling rings off and on
      winding watch
      rattling keys or change in pockets
      opening and closing books, boxes, drawers
      scribbling while telephoning or interviewing
      grasping objects, paper, handkerchief while talking or singing
      stroking self or objects
      tapping or drumming with fingers on body or table
      clenching and unclenching fists, twiddling thumbs
      hitching up clothes
   3. Movements involving hands and face
      biting nails, scratching head or nose with fingers, pipe stem or pencil
      sucking pen, pencil, or finger
      picking nose, rubbing eyes
      pulling and twisting hair, ears
      running hands through hair, smoothing it back

[1] Cf. Krout (*93*), Olson (*124*), Freud (*55*).

adjusting spectacles or glasses
manner of smoking

4. Facial movements [1]

wrinkling nose or forehead, grimacing
squinting, rolling eyes, blinking, winking [2]
clenching or grinding teeth
rolling of tongue in mouth [3]
biting lips, licking lips, chewing gum
shaking, nodding head
humming, whistling, hissing
sniffing, coughing, sighing, clearing throat
yawning when not fatigued

5. Larger synkinetic movements of limb and body

flexing and extending arms or legs
crossing and uncrossing legs or feet
contractions of toes or other muscles not observable
shrugging shoulders
fidgeting and contortions of shoulders and limbs
leaning or pressing against wall or table
tapping feet

F. Pathological [4] (almost any of the above movements in I E may reach the stage of tics and compulsions)

tremors, choreas
organic spasms, cramps, convulsions, ataxias
verbigeration, echolalia, echopraxia
astasia-abasia, atonia, hypertonia
etc.

II. Standing, Walking, and Related Activities

A. Standing

1. Arms and hands [5]

(a) Symmetrical

hanging loosely
thrown back
elbows bent
crossed in front, arms interlocked

[1] Cf. Lersch (*99*).
[2] Cf. Ponder and Kennedy (*132*).
[3] Cf. Scheck (*158*).
[4] Cf. Morgan (*144*).
[5] Cf. Roback (*141*).

hands holding upper arms or elbows
hands clasped in front or behind
hands in pockets, relaxed or tense
palms or backs of hands on hips
holding back of chair or other object

(b) Asymmetrical
one hand in pocket, on hip, holding lapel or some object

2. Fingers
closed in fist, or half bent
held together or apart
thumb pressed against forefinger or separated
little finger or forefinger pressed against others
thumb enclosed in other fingers
various fingers crossed

3. Legs
standing evenly on both feet
standing on one leg, other relaxed or placed on some higher object
frequent shifting from one foot to another
one foot behind the other
legs pressed together or wide apart

B. Walking [1]

1. General qualitative characteristics
items under I A 1

bobbing	plodding
mincing	precipitate
nautical	tripping

2. General quantitative features [2]
length of stride
speed of walk
definiteness of direction
elasticity
weight and noise
regularity and changeability
rhythm; influence of counter rhythms (*e.g.*, walking with a second person, or during band music)

[1] Cf. Bogen and Lipmann (*18*).

[2] Cf. Wilsing (*18*).

raising of feet or shuffling
flat-footed, on toes, or heel and toe

3. Idiosyncrasies [1]

body upright, bent
arms swinging strongly, weakly or not at all
shoulders or trunk swaying
preference for walking stick
regular or irregular spirals when walking blindfold [2]
manner of using stairs

C. Strolling

items under II B
contrasts between strolling and walking habits

D. Running

items under II B
contrasts between running and walking

E. Athletics

items in I A 1
special analyses possible for quantitative and qualitative features manifest in

aviation [3]
driving a car
horseback riding
rowing
tennis
running
boxing

F. Dancing

1. Style and type [4]

items in I A 1
frequency of dancing
preference for old-fashioned, modern, folk, exhibition, or solo dancing
preference for simple rhythms, complicated rhythms
preference for closed position, open position
formal and stiff, or experimental and supple
adaptable (from partner to partner)

[1] Cf. Galton (*56*).
[2] Cf. Schaeffer (*157*).
[3] Cf. Morf's discussion of the patterning in loops and flying stunts of aviators (*111*).
[4] Cf. Flach (*49*).

strong or weak as leader (man)
submissive or resistant as follower (woman)

2. Rhythmic factors

accurate sense of rhythm or little sense
rhythm carried in whole body, in legs only
synkinesis of trunk, arms, hands
preference for quick time, slow time
preference for syncopation, for gliding
preference for long steps, short steps

G. Performance of Routine Tasks

items in I A 1
special analyses possible for quantitative and qualitative features manifest in

dressing and undressing [1]	eating
washing and shaving	household chores [2]

III. Sitting and Resting

A. Sitting in a Chair

items in I A 1, I B, I E

1. Arms and hands

hands clasped on lap, across stomach
hands resting in lap, clasping knees
hands clasped behind neck or on top of head
one hand resting in other, or fingers interlocked
resting arms on chair or other support
hands supporting chin or cheek
palms or backs of hands against face
elbows supported on legs, table, or arms of chair

2. Legs and feet

one or both legs straight or bent
legs crossed; uppermost leg swinging or still
feet crossed
legs twisted around one another
legs wide apart or close together
feet drawn under body
feet twisted around chair or resting on rungs

[1] Cf. Wagoner and Armstrong (*190*). [2] Cf. literature on Taylorism.

preference for high seats, dangling legs
resting feet on footstool, chair, mantelpiece

B. Lying Down

special analyses possible for quantitative and qualitative features manifest in

resting	sleeping [1]

C. Sedentary Occupation

special analyses possible for quantitative and qualitative features manifest in

reading	working at a machine
writing	sewing
typing	playing card games
playing musical instrument	playing chess, checkers

IV. Communication and *Mimik*

A. Talking

1. Attributes of voice [2]

soft or loud
dull, sharp, shrill, nasal, or hoarse
smooth, melodious, sonorous
rapid, moderate, or slow
constant or variable
changes: swift, gradual, periodic
dropping voice at end of sentence
rising inflection throughout sentence
uneven accentuation of words, staccato
even accentuation of words, legato

2. Style or type of speech

loquacious, taciturn
frequent or infrequent talking
circumstantial or meager
clear or confused
rich or poor in vocabulary
favorite words and expressions
aphasias, pauses, stammering, etc.

[1] Cf. Johnson and Swan (*80*).

[2] Cf. Sapir (*147*) and Pear (*130*).

lisping or other mannerisms
long or short sentences
complete or clipped phraseology
number of interjections
style of retelling [1]

3. Motor attitudes during conversation

much or little synkinesis
imitative and sympathetic response to speaker
expressionless attention
inattention or impatience
following speaker with eyes

B. Laughing

1. Chiefly qualitative attributes

penetrating	soft, rippling, melodic
explosive	wooden, flat
hearty	infectious
cackling	grunting
chuckling	

2. Chiefly quantitative

loud or soft	whole body, face and voice only
frequent or infrequent	noiseless or noisy
intermittent or continuous	

C. Weeping

frequent or infrequent	with or without tears
conspicuous or inconspicuous	audible or inaudible
	shrill or quiet
with whole or part of face	quivering
with pouting	

D. Gesticulation

items in I A 1

frequent or infrequent	use of trunk
use of one arm or both	to emphasize speech or when words fail

[1] Cf. Wolff (*199, 200*).

E. Handshake

1. Pressure

strong or weak
long or short duration
single or repeated
swinging, clasping
active (with initiative), passive

2. Manner of offering the hand

open, or fingers semi-closed
palm up or down
arm high or low
elbow straight or bent

3. Patterns of handshaking

assured or timid
with compensatory vigor
using both hands
grasping other's shoulder
patting other's back
kissing hand or raising it to breast (cavalier style)

V. Handwriting [1]

speed: average and ratio of normal to maximum
average pressure (point and grip)
rhythm of pressure
area of writing
size of normal small letters, of capitals
spacing of lines, words, margins
breadth of letters
length of upper and lower projections
alignments: rising, falling, sinuous or straight
slant of letters: average, at beginning, at end
variability in breadth, slant, alignment
connectedness and brokenness
manner of connections: garland, arcade, threadlike, pasty
simplification and embellishment
Form niveau: innumerable patterns

VI. Other Accessible Fields of Expression

A. Disposition of Time: regularity of habits, punctuality in keeping appointments; amount of wasted time

[1] Cf. Downey (*39*), Enke (*48*), Klages (*85*, *86*, *87*), Krauss (*92*), Saudek (*148–155*), Ziehen (*205*).

B. Dress: color, neatness, fashion, expense; style and quantity of ornament, use of scent

C. Furnishings: style, type, and taste in room or house furnishings, automobile accessories

D. Literary Style: diary, correspondence, compositions

E. Visual Arts[1]: manner, if gifted, of drawing, painting, etching, photography

F. Musical Arts: style of composition[2] and performance

G. Manner of Appreciation: styles preferred, and way of responding when pleased or displeased

H. Public Appearance: change or naturalness in bearing, nervousness, self-confidence

I. Scientific Work: mechanical construction,[3] architecture, drafting, manner of research

J. Games and Recreation: amount and manner of play; behavior during psychological tests[4]

K. Nickname: often signifies prominent features of expression as apprehended by others (*e.g.*, "Lightning," "Weeping Willy," "Dizzy," or "The General").

[1] Cf. Lewis (*102*) and Prinzhorn (*136*).
[2] Cf. Roback (*142*) and Vernon (*186*).
[3] Cf. F. H. Allport (*2* p. 107).
[4] Cf. Burt (*26*).

CHAPTER II

METHODOLOGICAL CONSIDERATIONS

The Experimental Sessions. The main part of the investigation consisted of three experimental sessions, all individual, with 25 subjects. Between the sessions there was an average interval of about four weeks, so that samples of each subject's performances were obtained over a period of approximately eleven weeks. In the discussion of any experiment, the session at which it was performed will be indicated by Roman numerals (I, II, or III).

Each experimental session was first tried out with a few volunteer subjects not included in the regular group of 25. The experiments proper were conducted by one of the writers, usually with the aid of an assistant, in order that double records might be obtained. Since much of the recording had to be done as unobtrusively as possible, often under difficult conditions, this precaution was highly desirable. The discrepancies between the two sets of measurements, however, were seldom as great as 5%, so that the two records were always averaged unless one was known to be much more dependable than the other. During session II when almost all the observations were recorded kymographically, a second experimenter was unnecessary.

In all sessions the experimenters observed the subjects and made notes on their behavior. Normal conversation and activity were encouraged in order to bring out spontaneous motor characteristics. Some aspects of expressive movement which could not be tested were studied

by means of rating sheets (pp. 93–95) filled out by each experimenter after sessions I and III. After session III each subject filled out the same sheets with self-ratings, and obtained ratings from several friends and associates.

Before the actual tests and experiments can be described (in Chapters III and IV), and before the treatment of the results can be understood (Chapters V and VI), it is necessary to discuss some of the conditions and assumptions of the investigation.

The Subjects. The subjects, all of whom were male, ranged in age from 18 to 50 years, and were drawn from diverse backgrounds. Seven were undergraduate volunteers, nine were somewhat older graduate students at a theological school, and the remaining nine were older business or professional men, volunteers from a university extension course in psychology. These subjects were selected intentionally to secure a reasonable degree of diversity. It is well known that heterogeneity directly affects the size of every coefficient of reliability or correlation. It therefore seemed desirable to aim at a more representative population than that usually obtained in the highly homogeneous groups of student subjects. The best indication of the actual amount of heterogeneity is, of course, the range of individual differences obtained. In Chapter IV, therefore, the extreme measurements and the median (*i.e.*, the 13th highest score) are given for each experiment.

The Order of the Experiments. To offset the influence of mood and other temporary determinants of the subjects' performances, most of the tests were administered twice, and at different sessions. In some few cases the repetition occurred in the same hour; it was then separated from the original test by as large an interval of time as the session would permit. The final measure

on each test was derived from the sum or from the average of these two sets of observations. The reliability of the measure was, therefore, determined by the correlation of the subjects' rank orders in the two series, corrected by the Spearman-Brown formula.

It was necessary likewise to control the influence of mental set, space and time errors, and perseveration, in the performance of all tasks involving similar psychological functions. In order to do so, all similar tasks were separated as much as possible during a given session. For example, at session I, a test of the estimation of distances with both right and left hand was applied; but between the two experiments there were a number of entirely different tests such as blackboard writing and drawing, estimation of weights, counting, and arranging Binet cubes. In Chapter IV the various experiments are described in detail according to a convenient scheme, *not* in the order in which they were carried out.

Errors of Sampling. For reasons to be explained below, no elaborate statistical analysis was attempted. The original measurements (times, distances, etc.) were almost always transmuted into rank orders, and the rank order method of correlation was applied.

Twenty-five subjects are, of course, a far from adequate sample; from the statistical point of view the probable errors of sampling will be large. In order to avoid continual citation of probable errors, the significance of various degrees of correlation will be given once and for all in the following table.[1] This should be read:

[1] Rank order coefficients (ρ's) are often corrected by a small amount to give the probable value of the equivalent product moment coefficient (r). In the present study, all coefficients have been left as ρ's, but the indices of probability in this list have been calculated to be equivalent to those that would be obtained had the corrections been applied.

a ρ which is equal to or greater than $\pm.13$ is at least 1 times its probable error, and so forth.

ρ	$\frac{\rho}{\text{P.E.}}$
$\pm.13$	1
.25	2
.30	$2\frac{1}{2}$
.35	3
.43	4
.50	5
.69	10

The major portion of the coefficients mentioned in our results do not attain a high degree of statistical significance. In most work on individual differences, this might be a fatal objection, since investigators usually set out to find the exact relationship between two definite and specified variables. But the aim of the present study is somewhat different. It is the *general run of correlations* between a wide range of different variables which is of primary interest here. Our conclusions seldom depend on any single coefficient, but when they do, they are based on coefficients which are at least three times their probable errors.

This dividing line of "statistical significance" (at three times the P.E.) is, of course, an artificial one. A coefficient of +.36 which lies above the border is not greatly superior to +.34 which lies below. A person whose age is 21 years and 1 day is not always a man, nor is one whose age is 20 years and 364 days always a boy. But, like the 21st birthday, a coefficient which is three times its P.E. serves as a conventional and convenient boundary.

Correlations between Composites of Tests. In dealing in this study with a large number of measures of the same variable, or combinations of variables, three simple

statistical methods will be of considerable use. (1) In determining, let us say, the reliability of ratings of 25 subjects by eight judges, one might sum the ratings given by four judges and correlate these with the sum of the ratings by the other four. It may be shown, however, that the extent of agreement among all the judges is a direct function of the relation between the standard deviation or dispersion of the original separate ratings (σ) and the S.D. of the final averaged ratings (σ_a). The formula for the average intercorrelation between all the ratings, $\bar{r}_{11}$, is given by Kelley (*82* Formula No. 171), as follows:

$$\bar{r}_{11} = \frac{\frac{\sigma_a^2}{\sigma^2} - a}{a^2 - a}.$$

Here a is the number of raters, or the number of series of observations that are combined. The reliability of the sum of the ratings is equal to this average intercorrelation stepped up a times, *i.e.*,

$$r = \frac{a\left(1 - \frac{a\sigma^2}{\sigma_a^2}\right)}{a - 1}.$$

The same method will be applied in computing the reliability of a large number of observations of muscular tension (p. 92).

(2) If instead of combining original measurements, we wish to combine rank orders, the same formula may be adapted as follows (Kelley No. 172):

$$\bar{r}_{11} = 1 - \frac{a(4N + 2)}{(a - 1)(N - 1)} + \frac{12\Sigma S^2}{a(a - 1)N(N^2 - 1)}.$$

Here S is the sum of the ranks for any one of the N subjects on the a tests.

(3) Finally, if we wish to compute the correlations between one composite of tests $1 + 2 + 3 + 4 + \cdots + a$, and a second composite $\mathrm{I} + \mathrm{II} + \mathrm{III} + \cdots + b$, knowing the individual correlations between all the separate tests, we may apply Spearman's method of sums (*167*).

$$r(1 + 2 + 3 + \cdots + \mathrm{a})\,(\mathrm{I} + \mathrm{II} + \mathrm{III} + \cdots + b) = \frac{\Sigma r_{1\mathrm{I}}}{\sqrt{a + 2\Sigma r_{11}}\ \sqrt{b + 2\Sigma r_{\mathrm{II}}}}$$

Σr_{11} is the sum of all the correlations between tests of the type 1, 2, 3, $\cdots$ a.

Σr_{II} is the sum of all the correlations between tests of the type I, II, III, $\cdots$ b.

$\Sigma r_{1\mathrm{I}}$ is the sum of all the correlations between pairs of tests such as 1 I, 1 II, $\cdots$ 1b, 2 II, $\cdots$ ab.

By making a equal to unity, the same formula gives us the correlation between test 1 and the sum of tests I + II + III + $\cdots$ + b, if we know the individual correlations $r_{1\mathrm{I}}$, $r_{1\mathrm{II}}$, $\cdots$ r_{1b}.

Definitions. It will be well to define here certain terms which will be used throughout the following chapters to describe several orders or planes of combination and intercorrelation.

Raw scores = the original measurements, expressed as times, distances, areas, weights, or ratings by individual raters. Roughly 300 sets of such raw scores were obtained from each of the 25 subjects.

Original measures = rank orders based on single sets of raw scores, derived from the application of one experiment at one session.

Paired measures = rank orders based on the averages

of two sets of observations obtained under identical conditions, or the combination of two original measures, one of which is an exact repetition of the other.

Combined measures = rank orders based on the sum of two, three, or more similar original measures which are not exact repetitions of one another. Averaged ratings on a single quality also constitute a combined measure. The 300 original measures give roughly 100 paired or combined measures in all.

Composite measures are derived from a number of more or less unlike original paired and combined measures which agree together rather closely.

Variables are the 40 or so composite, combined, or paired measures to which original measures are reduced. (Cf. *Appendix.*)

"*Group factors*" are obtained by combining half a dozen or more variables. As will be explained below, this term was adopted as being the most convenient, without, however, implying the statistical connotation which Spearman has given to it. The half dozen possible group factors discussed in Chapter V are the highest level of association discovered in the present investigation. They practically correspond to what might be called "psychomotor traits" of personality.

Repeat reliability applies to the correlation between two identically obtained original measures. *Corrected reliability* applies to the corrected average intercorrelation of two or more components of a variable; it may be used interchangeably with the term "internal consistency."

Internal consistency is the more usual term for the corrected reliability of broad variables, *i.e.*, of composite measures or "group factors." If we find that a certain composite variable has a corrected internal consistency of +.80, derived from the average intercorrelation of,

say, 5 component variables, we mean that +.80 is the theoretical or probable degree of correlation between our five measures, summed in one composite rank order, and another similar composite of five more like measures.

Principles of Building Combined and Composite Measures. Examples of the derivation of these aggregates or associates of measures will clarify their meanings. Let us start with normal speed of tapping with the forefinger. This performance was measured with the right and left fingers at session II, and again with the right finger at session III. The correlations between the three *original* measures were:

R. finger II, III = +.86; R. finger II, L. finger II = +.81; R. finger III, L. finger II = +.91.

Thus the repeat reliability of the *paired* measure for speed of tapping with the right forefinger is +.86, corrected +.93. But in this, and in many similar experiments, it was found that the correlations between the performances on different sides of the body were as high as the correlations between repeated performances with identical muscle groups. In such cases all three original measures are summed to give a *combined* measure. The average intercorrelation of these three speeds of finger tapping is +.86, giving a corrected reliability for our combined measure of +.95.

Composite measures are still more general. They were isolated on the basis of two considerations, reliability (or internal consistency) and independent variation. If the original measures failed to agree sufficiently highly when only two of them were summed as a paired measure, or three or four as a combined measure, several similar sets of observations had to be included in order to attain satisfactory corrected reliability. For example seven performances involved the grip pressure of fingers on a stylus

(in writing, tapping, etc.). (Cf. pp. 91–92.) Two of these performances were exact repetitions, at a later session, of two others. The corrected repeat reliabilities of these paired measures were only +.27 and +.50. Again two of the performances were recorded for the right hand twice and the left hand once. The corrected reliabilities of these combined measures were +.42 and +.59. But when all seven are compounded, we get an average intercorrelation of +.233, and an internal consistency, or corrected reliability of +.68. Since this is the most highly reliable variable that could be extracted from these performances, none of the lower orders of combined or paired measures were treated independently any further.

On the other hand, some of the paired or combined measures might be highly reliable in themselves, but they correlated so highly with one another, that they could not be taken as discrete variables. For instance, tapping speed was recorded for fingers, hand, and legs. But the correlations all ranged between +.78 and +.91, so that the muscle groups could hardly be considered separately. The unlike performances often correlated more highly than the like ones. Thus when the six original measures are compounded (two with right forefinger, one with left forefinger, one with right hand, and one with each leg), we get an average intercorrelation of +.865, and a corrected reliability for the composite measure of +.97.

In another set of experiments, subjects were told to draw circles the estimated size of 25¢ and 50¢ coins, and rectangles the estimated size of dollar bills. The corrected repeat reliabilities for the overestimation of the three sizes were +.74, +.83, and +.78, respectively. But the correlation between the paired measures for the 25¢ and 50¢ coins was +.87, so that it would be impossible to treat these two sets of estimations as separate variables. On

the other hand, the correlation of Overestimation of Dollar Bills with the combined or composite measure for Overestimation of Coins was only +.34. These two variables were therefore regarded as relatively independent or discrete, and *not* added into a broader composite consisting of overestimation of coins *and* bills.

Corrections: Spearman-Brown Formula and Attenuation. The technique of building up more general combined or composite measures from a large number of interrelated original measures entails the frequent application of the Spearman-Brown prophecy formula. Now with a small number of subjects, and with data whose reliability, though rather satisfactory, is far from perfect, one cannot be too cautious in the application of statistical "manipulation." The writers have accepted the use of corrections for reliabilities and internal consistencies, chiefly as a convention, though at the same time such corrections do not seem to them to advance in any way the *psychological* significance of the results. It should not be forgotten also that the Spearman-Brown formula implies certain statistical conditions which are not always fulfilled by data of this type.[1] For these reasons the uncorrected, raw figures are always given in the text, and it is on these that the writers have attempted to base most of their conclusions.

Let us take an example which is an extreme and yet logical extension of the methods of correction frequently applied. It will be shown below that 24 of the final variables possess an average intercorrelation of +.055. This fact may be explained by the existence of some general

[1] The formula is a special case of Spearman's earlier method for correlation of sums; cf. above, p. 41. Actually it serves to predict the correlation between the sum of a number of tests 1, 2, 3, etc., and a hypothetical battery of similar tests I, II, III, etc. The simple Spearman-Brown formula is applicable *only* when the average intercorrelation between 1 2, 1 3, 2 3, etc., is equal to the average intercorrelation between I II, I III, II III, etc., and equal to the average intercorrelation between 1 I, 1 II, 2 I, 2 II, etc. In work of this type, the condition is seldom likely to be strictly achieved.

"psychomotility" or "vitality" factor which is present in all the tests; or it may be due to chance (cf. pp. 105–108). If we step up +.055 twenty-four times, we obtain an internal consistency of +.58, that is, a theoretical validity of +.77, representing the probable correlation of the sum of the 24 variables with an infinite number of other similar variables. So that, *theoretically*, this sum of the 24 variables actually measures a psychomotility factor with a considerable degree of validity. Moreover the variables themselves possess independent reliabilities of but little over +.80, on the average, so that if our general factor were corrected for attenuation, it might well attain a much higher validity. Nor is this as absurd as it appears at first sight. The 24 variables were divided at random into two groups of 12, and the correlation between the two summed halves was found. This more empirical method exactly fulfilled the statistical prediction, since the reliability of the 24 variables proved to be +.59.[1] And yet in a case like this, it seems to the writers more psychologically sound to admit that their variables possess an almost negligible average intercorrelation of +.055, rather than to claim that they have measured a general factor with a validity of near .90.

Corrections for attenuation are not employed in the treatment of our results since they imply certain statistical assumptions which are seldom fulfilled among small groups of subjects [2] and since, in a study such as the present one, the corrections add nothing to the psychological significance of the results. If the composite variables or the group factors, as reported below, are substan-

[1] Cf. also Ruch, Ackerson, and Jackson (*145*), who showed empirically that, with homogeneous test material, the predictions of the formula are confirmed with a considerable degree of accuracy.

[2] *E.g.*, that the chance errors of measurement, eliminated by the attenuation correction, are themselves independent and uncorrelated.

tiated by further research on a larger sampling of the population, then the more elaborate methods, including not only corrections for attenuation but also partial and multiple correlation techniques and differential weighting, would undoubtedly aid in the interpretation of the data. The writers, however, intentionally used but few subjects and a large variety of tests, in an attempt to map out a broad and almost unexplored territory in a tentative way.

Intrinsic Inadequacies in the Statistical Treatment of Consistency. When quantitative methods are applied to the complex problems of the organization of personality they uncover only partially, and sometimes not at all, the consistencies and interrelations which naïve common sense leads us to expect. This is a familiar phenomenon, manifested to some extent in the results of the present study, as well as in virtually all of the "exact" research in personality to date. This situation may be due partly to the tendency of common sense to employ clichés, to oversimplify its impressions of personality, accenting dominant traits and disregarding complexity, with the result that it ascribes homogeneity, integration, and consistency to an individual to whom they do not rightly belong. But, on the other hand, the fault does not lie entirely on the side of common sense. There are certain inherent limitations in the statistical method which inevitably prevent it from revealing the full unity of personality. These should be briefly considered.

(1) Psychometrics cannot deal directly with single individuals; it has to approach them indirectly through standards established for a representative group. In the present investigation, for example, the relations between the various performances of any one subject can be determined only by comparing his rank order or percentile

standing on one test with his standing relative to the group on each of the other tests. The *patterning* of characteristics within the single individual is entirely inaccessible to direct quantification.

(2) Exact experimentation is forced to oversimplify, to select some specific aspect of behavior on which all subjects can be quantitatively compared. This procedure results in the neglect of innumerable factors, different in each individual, which determine the particular bit of behavior under examination. In the attempt to establish one significant co-variation which will be valid for the whole group, these important individual determinants have to be averaged out and blurred. In short, the endeavor to compare what is tangible and measurable may lead to a comparison only of what is coarse and uncomplicated. Here again *personal* consistency is disregarded by the exigencies of method.

(3) When supposedly related measures fail to correlate to the degree that convention accepts as significant, the statistical method may create a false impression that the problem of consistency in personality is solved in the negative; that specificity reigns supreme. The truth may be that even low correlations may be due not to chance factors (as is usually supposed), but to the conflict of several highly consistent dispositions. For instance, in our motor case study No. 1 (pp. 134–137), we find a relatively rapid natural walking tempo but a very slow speed of strolling. These two records for this single subject lower considerably the statistical consistency of our measures of speed. This case study will show, however, that this subject's speed of walking is associated with a "trait" of aggressive and assured behavior, whereas his speed of strolling is a reflection of an equally well-organized "trait" of thoughtfulness and

caution. Our subject, then, has genuinely consistent dispositions which are of a rather individualistic order and so operate to reduce the size of our statistical measures of consistency in the group as a whole. Throwing together a subject whose consistency lies in an unexpected direction with others whose consistency lies in the expected direction lowers our statistical measure, without in the slightest degree disproving the hypothesis that personality is highly consistent. If the statistical method is to be used at all, this consideration should dispose us to regard low but positive coefficients as more valuable indicators than conventional practice allows.

(4) Finally it should be remembered that the origin, rationale, and psychological meaning of numerical results lie quite outside the scope of the statistical method. Thurstone writes, "I do not believe that these correlation methods and particularly the reliability formulae have been responsible for much that can be called fundamental, important, or significant in psychology. On the contrary, the correlational methods have probably stifled scientific imagination as often as they have been of service. As tools in their proper place they are useful but as the central theme of mental measurement they are rather sterile" (*183*).

Other limitations of statistical procedure have been set forth by Myers (*119*), Wilson (*196*), and the Thomases (*180*). But the considerations we have listed will be sufficient to warn the reader that while applying conventional quantitative methods we regard them only as part of the foundation upon which to rest our conclusions. The reader, therefore, should not feel oppressed by the preponderance of statistical matter in the ensuing chapters. In every case the psychological methods of direct experimentation, descriptive case study, and critical

interpretation, take precedence. It is hoped that this explanation together with the frequent critical summaries, statements of conclusions, and interpretations, throughout the remaining chapters, will reassure the reader who distrusts excessive quantification in studies of personality.

CHAPTER III

TECHNIQUES FOR RECORDING MUSCULAR PRESSURES AND TENSION

Many of the experiments conducted in the present investigation required only the simplest apparatus—such as a stop watch and a ruler. Handwriting, tapping, and other activities that involve muscular tension and pressure, however, demanded more elaborate methods of recording. Before proceeding to the experiments themselves, therefore, it is desirable to review the techniques devised by previous investigators in this field, and to describe the apparatus constructed for the present study.

We may distinguish three kinds of pressure and tension. In writing or tapping there is first the pressure exerted between the point of the pen, the stylus or the finger and the paper or tapping surface. Experimental graphologists (*vide* Saudek *152*) call this *point pressure*. Secondly there is the pressure with which the fingers hold the pen or stylus in writing, tapping, or tracing mazes. This is known as *grip pressure*. Thirdly there are the accompanying tensions in other muscle groups during such activities as writing or tapping, as well as general muscle tonus when the subject is resting passively. There are several available techniques for the study of these various aspects of pressure and muscular tension.

Point Pressure. (1) *Microscopic Inspection.* Fineness or heaviness of line in a script is, of course, directly dependent on point pressure. As Saudek's (*150*) and Osborn's (*126*) facsimile reproductions show, it is possible to determine fairly exactly, when the script is

considerably magnified, the amount of pressure exerted on the point of the pen at any stage in the writing. This method is obviously unsatisfactory, however, since the width of the stroke depends likewise to some extent on the softness or hardness of the pen point, or the thickness of the pencil point. The distribution of the pressure may be roughly determined from the relative thickness of the strokes in a single script written with a particular pen, but it is rarely possible by this method to determine the absolute, average point pressure. Delhougne (*34*) has overcome some of these difficulties by making the writers use a specially constructed pen with three points instead of the usual two. The splaying of these points, when inspected microscopically, gives a reliable record of the pressure.

(2) *Carbon Paper Method.* Bills' (*12*) and Baxter's (*8*) *carbon paper* technique is almost as simple as these observations of width of line, and permits of much more reliable quantification. Several sheets of thin paper are interleaved with carbon paper, the whole set of sheets being firmly attached to a writing board. By examination of the sheets successively lower down in the pile, it is possible to grade the degree of pressure. Bills and Brown (*14*) set their subjects varying numbers of arithmetical problems to work out, and investigated the effect of the number of problems to be done upon the efficiency of performance. The strength of this quantitative set, they found, was closely related to the muscular tension revealed by the depth of the carbon paper impressions. The errors of measurement of the pressures were found to be small. Two judges independently scored the sheets, and agreed with one another as to the amount of pressure exerted by each subject during each experiment to an extent usually greater than a correlation of +.90.

Baxter (*8*) employed this method for measuring handwriting pressure as a possible indicator of "temperamental strength." While she obtained a repeat reliability of $+.833 \pm .03$, there was little agreement with her other diverse measures such as blood pressure, dynamometer strength, contra-suggestibility, and the like.

The chief advantage of the method is that the subject is not distracted by an unwieldy stylus, nor by a writing surface which "gives." In Baxter's work, no subject suspected the presence of the carbon sheets. Moreover the mere weight of the hand on the paper does not affect the impressions, and this is a difficult factor to control in the more commonly employed tambour or balance methods (see below).

In the present investigation, the method was as follows. Before the session (I) started, ten sheets of thin paper were arranged alternately with ten sheets of thin carbon. On top was placed a single opaque white sheet, slightly larger in size than the rest. When the whole set was pinned to a drawing board, the carbon and other lower sheets were hidden; only one subject noted their presence. A newly sharpened pencil was used by each subject in writing. When the depth of penetration was estimated, later, the lower sheets were all examined by daylight. As the impressions, from the top downwards, gradually disappeared, the five most clearly marked lines or letters were selected, and followed through until each one was no longer visible. The score was the sum of the depths of penetrations of these five letters, or parts of letters. Only two subjects penetrated more deeply than the total 21 sheets. The written material included the subject's signature; since this was often accompanied by a consistently greater pressure than the rest, not more than two out of the five impressions were ever taken from the signature.

(3) *The Kraepelin Balance.* The principle of the *Kraepelin balance* has been applied in the majority of investigations of point pressure. The method is analogous to writing on some kind of weighing machine, from which the weight as shown by the pointer on the scale can be read at every instant. A revolving smoked drum is usually substituted for the scale, so that the lever which is raised by depressions of the writing surface, traces a continuous kymographic record. There are two obvious disadvantages to this technique. First, people are used to writing on a surface which does not yield under pressure; completely natural writing will therefore be difficult to obtain. This factor can be controlled to some extent, by using strong springs and very large magnification, so that the actual amount of yield of the surface is scarcely noticed. Secondly, the subject will desire to rest varying amounts of his hand on the writing surface, so that the tracing will be a mixture of true point pressure and hand pressure. If the writing surface is small, however, it is often possible to arrange a support for the hand, allowing only the pen point to touch the yielding surface.

The pioneer pedagogical work of Meumann (*108*) and F. N. Freeman (*52*) was performed with an apparatus of this type. In addition, Freeman inserted a typewriter ribbon and a constantly moving band of paper underneath the actual writing paper so that not only the pressure but the speed of writing was known for every detail of every letter. Enke's (*48*) investigation of the rhythms of writing among different constitutional types (cf. Chapter I), also employed the Kraepelin balance.

In the present study, the apparatus was constructed as follows. In the middle of a horizontal table, the usual height from the ground, was cut a rectangular hole. A board measuring 8″ × 6½″ fitted in the hole closely, but

without touching the sides. This board was fixed below to one end of a long horizontal, solid brass lever. The lever was pivoted on a horizontal axis some 18″ from the center of the board. Thus, although the board moved in a slightly circular path about this pivot, yet most of the ordinary deflections below the level of the surrounding framework (*i.e.*, the table) were practically vertical. The board was held, when under no pressure from above, on a level with the table by two coiled springs, whose strength had been adjusted to a suitable amount. In order to prevent rebounds of the board above the level of the table after a rapid depression, the board was tied below by a cord which did not hamper its downward movements. In the illustration of tapping and writing records (Figure I), the small extent of the rebound between taps, or written words, will be clearly seen. The movements of the further end of the brass lever were magnified and recorded on a smoked drum in the usual way. The whole apparatus was secured to the wall, and supported on a concrete floor, so that a high degree of rigidity and stability was obtained.

In writing or tapping, no attempt was made to conceal the "give" of the board from the subjects. Before writing any words, however, they had some practice in simpler tasks, and were thus able to approximate their normal manner of writing and their usual amount of pressure. To prevent the complication of point pressure with hand pressure, a thin rigid rest was laid across the right-hand bottom corner of the board, supported by the framework. The subject placed the ball of his hand on this rest, and wrote on the board diagonally. That this method was successful is shown by the kymograph records, for the lever returned to approximately the same zero base line between each word that was written. Thus the obtained curves must have been due to the point pressures of the

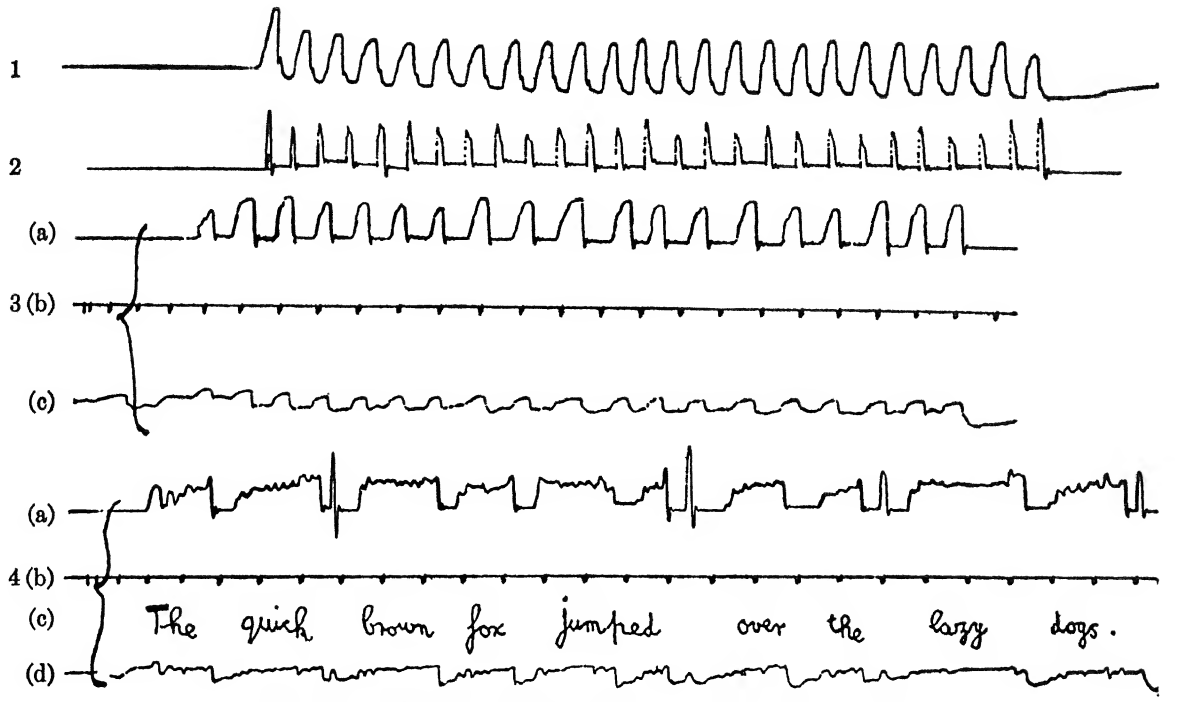


FIGURE I. NEGATIVE REPRODUCTIONS OF KYMOGRAPH RECORDS.

pencil or stylus on the board, not of the hand. Further, before the writing experiments started, the subject was told to lay the whole hand on the board, as if he were about to write in the top left-hand corner. This was repeated several times, to give an average measure of the pressure of resting hand. The correlation found between this hand pressure and the average writing point pressure was $-.02$.

Since rank orders for pressure and not absolute measurements were required, the exact pressure for each deflection of the kymograph lever need not concern us. It was found by calibration, however, that with most ordinary ranges of pressures, the board would sink approximately 1 mm. per 100 gms., and that the lever would rise 5 mm. Thus in the illustration of a finger tapping record (Figure I, 2), where the average height of the curves is about 1 cm., we may deduce that an average pressure of 200 gms. was exerted by the finger of the tapper.

(4) *The Tambour Method.* This method is identical in principle with the Kraepelin balance, except that an air

---

EXPLANATION OF FIGURE I.

1. Compression of stylus for 20 seconds between fingers of right hand (cf. pp. 79–81).
2. Tapping on pressure board with right forefinger for 20 seconds (pp. 79–81).
3. Drawing parallel lines on pressure board (pp. 85 f.).
   (a) Point pressure
   (b) Seconds
   (c) Grip pressure
4. Writing the words "The quick brown fox jumped over the lazy dogs" (pp. 87–91).
   (a) Point pressure
   (b) Seconds
   (c) The actual words written on the smoked drum afterwards by the same subject
   (d) Grip pressure

connection replaces a lever connection to the kymograph. Morgan (*113*) wished to find the effects of distraction on a continuous choice reaction test. His six tapping keys were arranged, unknown to the subjects, on a stiff tambour. The force of each tap depressed this tambour slightly, and the expelled air raised the lever of a second small tambour. The average pressure per reaction varied for six different subjects from 100 to 1120 grams, and it was invariably increased (usually in a ratio of about 3 to 2) under the influence of distractions.

Though the authors fail to describe their apparatus, it is probable that Smith, Culpin, and Farmer (*166*) used a similar method for recording tapping pressure during telegraph operations with Morse keys. They found that telegraphists who were liable to telegraphists' "cramp" exerted, on the average, nearly twice as much pressure as a control group of operators. A large tambour, whose "give" was not noticed by the subjects, was used by Ford (*51*) in investigating the effects of distraction on the written performance of arithmetical problems.

(5) *Stylus Methods.* Tambours of this type, of course, include the hand pressure along with their records of the point pressure of writing. To eliminate this factor entirely, Drever (*41*, *42*), constructed the first stylus instrument for measuring point pressures. In this stylus, the pen or pencil slid in an outer tube or holder which was grasped by the subject. To the top of the holder was fixed a small tambour, above the level of the subject's hand, and the actual pressure between the point of the pen and the paper resulted in depressions of this tambour by the inner sliding shaft. From the tambour a rubber tube connected with a kymograph.

More recently, Román-Goldzieher (*144*) has constructed a point pressure stylus along similar lines to that of Dre-

ver. The curves which she reproduces in her article, are very similar to those obtained by the present writers with their pressure board technique.

Bills (*12*) and Stroud (*177*) overcame the unwieldiness and the inaccuracies of Drever's stylus, by fixing the writing lead to a narrow rigid piston, sliding in the center of their stylus. This piston was connected at the end to a flexible coil wire, such as is used in kodak shutter releases. At the other end of the wire was attached a lever writing on a kymograph. Point pressures in writing simple figures, or in tracing a maze were thus recorded. Though the instrument is, at present, less delicate than the Kraepelin balance apparatus, it could no doubt be adapted to tapping, or to ordinary handwriting. The mechanical connection as opposed to the tambour connection with the kymograph is a considerable advance in technique.

Grip Pressure. For measuring grip pressure, the apparatus always takes the same form, that of a compressible stylus, connected by a rubber tube to a recording tambour. Drever (*41*, *42*) constructed a writing stylus (different from his point pressure recorder) with rubber bulbs at the points where the fingers grasped it. These bulbs and the connecting tube were filled with mercury, since the rubber was not sufficiently rigid for an air connection. But the instrument was therefore very heavy, and probably made the writing somewhat unnatural. Bills and Stroud combined their grip and pressure recording instrument in one stylus. Outside the sliding piston was a heavy rubber tube, held by the subject, from which a side tube led off at the top to a tambour.

Johnson's (*79*) tapping stylus was similar. It recorded grip pressure and its variations throughout a series of tapping experiments. Johnson concludes that "the differences in tension between individuals at the same level

of practice for the specific performance" (*i.e.*, maximum tapping speed) "suggest that this measurement may contribute to the study of individual differences in the fundamental tendencies often called temperamental traits." Among adult subjects Johnson noted four types of pressure curves during tapping:

(a) low constant pressure (the most accurate and quick performers)
(b) irregular but high average pressure (the least efficient)
(c) initial rise in pressure, then a gradual decrease
(d) slight pressure, gradually increasing throughout.

He believes that those who showed more irregular pressures are "easily disintegrated" personalities.

For recording grip pressure in tapping or writing in the present investigation, a similar stylus was constructed. In the center was an inner shaft consisting of a thin aluminum tube, along most of the length of which were cut narrow slits. Thus the shaft yielded freely to compression, but it could not be bent. The writing lead was attached to the bottom of the shaft. Outside the shaft was a fairly thick rubber tube, of larger diameter (roughly equal to that of a normal fountain pen), from the top of which a narrow, light rubber tube led off to a sensitive tambour. This connecting tube was supported above, so that it might impede the writing or tapping movements as little as possible. This apparatus was less unwieldy than either Drever's or Bills', being lighter than most fountain pens. Unfortunately it was not, however, equally compressible at all points. Though all the subjects were directed to grip the stylus at the same place, yet it was impossible to eliminate all the differences due to position, and the records obtained will be found to be of unsatisfactory reliability. Moreover, in spite of the great magnification.

the amount of compression was very small, and the records were not easy to measure. In the drawing record illustrated in Figure I, 3a represents the point pressure, 3c the grip pressure. The reader will see that the tracings of grip pressure seldom return to the same zero point, while in the point pressure tracings, the zero level is more clear-cut. This situation results in part from the fact that a lever connecting mechanism was employed in recording point pressures, and a tambour mechanism in recording grip pressures. Tambour connections in general seem to be less satisfactory than mechanical connections.

General Muscular Tonus. Most investigators have been interested in variations of tension shown by a single individual, and not in the average differences between individuals. G. L. Freeman (*53*) summarizes the methods for measuring alterations of tonus: knee jerk, electromyograms, thickening of the muscles as recorded by a tambour or a lever, etc. Jacobson (*78*) finds that all mental effort is accompanied by contractions of muscle groups (*e.g.*, visual imagery involves the activation of the ocular muscles). He believes that those who are more tense are mentally maladjusted, and that mental health is coincident with habitual relaxation of all unnecessary muscles. The most delicate work in this field is that of Golla and Antinovitch (*60*), who photograph the action currents in various muscle groups, finding that they usually increase in size during mental effort.

Bills (*12*, *13*) has developed coarser methods for recording tension in several muscle groups during mental work. Usually the subject holds a dynamometer at a natural pressure with his unengaged hand, or dynamometers with both hands when writing is not required. Alternatively, or simultaneously, sensitive springs are pressed against the sides of the jaw muscles. These are enclosed

in cones, and connected by a rubber tube to tambours. His results support the theory that increased muscular tension accompanies increased efficiency in almost any kind of learning. Moreover the recorded tensions are "increased by artificially inducing tension in other parts of the body." Dynamometer readings, jaw muscle records, and point and grip pressure tend to vary together rather closely. Tension, he finds, is quite consistent in all muscle groups of the body; this result is of considerable importance for our own investigation.

Duffy (*43*, *44*) suggests that this consistent muscular tension is "likely to find more general expression in young children than in adults." In adults the tensions may tend to be more specific, or differentiated among particular groups of muscles. For this reason she used 11 children (aged 3 to 5 years) in some interesting experiments on tension during choice reaction times under distraction. In the unengaged hand, the subject held a rubber bulb connected with a tambour and kymograph. Qualitative analysis of the tension curves showed high agreement with the children's personality traits, according to the clinical records of the institution where they were tested. A correlation of $+.52 \pm .15$ was found between average tension and ratings on excitability and emotionality.

Apparently the only attempt to measure actual individual differences in tension while the subjects are inactive is that of McKinley and Berkwitz (*121*). The most promising of their techniques was reproduced in the present study, and satisfactory measurements of the tonus in the flexor and extensor muscles of the elbow joint were obtained. McKinley and Berkwitz point out that the method could easily be modified for several other muscle groups.

The subject is seated in a chair and his lower forearm strapped to a horizontal board. Directly below the elbow

joint the board is pivoted on a vertical axis, so that the arm may rotate freely. The upper arm should make an angle of about 10 degrees with the horizontal. When the subjects are told to allow the arm to swing into the most comfortable position, it is usually found that the lower arm naturally assumes an angle of 90 degrees with the upper arm.

Below the board, on the same axis, is a large wooden pulley wheel. Horizontal strings lead from the circumference of this wheel to small brass pulleys at the same level, over which the strings hang. Hooks are attached to the ends of the strings, so that the application of a weight to either hook will immediately cause the wooden pulley, and therefore the arm, to revolve. A smoked drum is also attached to the wooden pulley and arm rest, revolving with the arm. The timing device is an ordinary 50 *d.v.* tuning fork, electrically maintained,[1] which vibrates in a vertical plane. In order to avoid vibrations of the arm rest, the tuning fork is fixed to a separate tripod; the vibrating lever may easily be moved up to the drum, as required. (This is our only important modification of the McKinley and Berkwitz design. They used a stationary smoked drum, while a time marker, electrically connected to the tuning fork, revolved with the arm and pulley.) When the arm has found its equilibrium, it is fixed by a stop, which can be released very quickly, so as to prevent rotation in either direction. The subject is told before each reading to relax his arm as completely as possible, that a weight will pull his arm outwards or inwards, quite harmlessly, and that he is to let the arm swing freely, without helping or stopping it in any way. His rôle is entirely passive and relaxed throughout the experiment. A weight is attached to the appropriate hook, the tuning fork is brought

[1] In many laboratories it should be possible to connect the time marker through a high resistance to the 50–60 cycle A.C. mains, thus eliminating the caprices of a tuning fork.

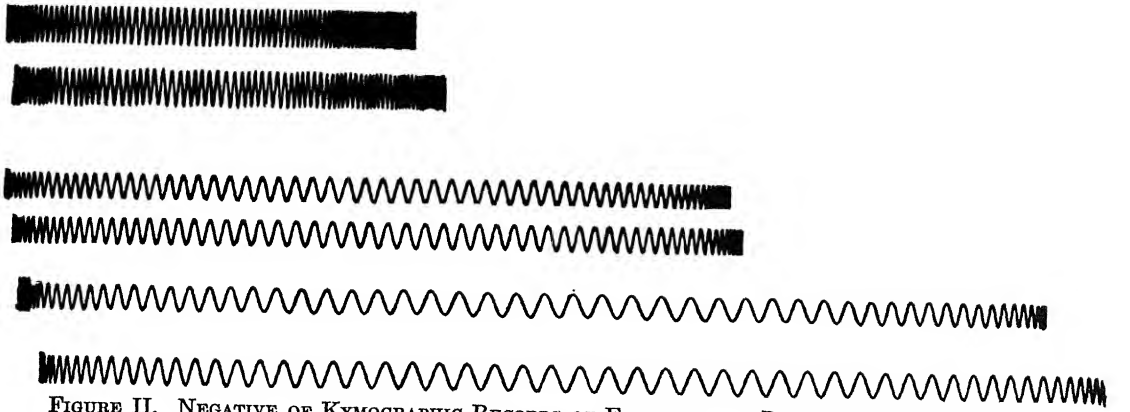


FIGURE II. NEGATIVE OF KYMOGRAPHIC RECORDS OF EXTENSION OF RIGHT ARM UNDER VARYING TORQUES (slightly reduced).

1 and 2: 515 grams; 3 and 4: 853 grams; 5 and 6: 1368 grams.

up to play on the drum, and the pulley is suddenly released. The arm starts swinging slowly, attains a maximum velocity, and is then gradually slowed up and stopped by the tonus of the muscles at the elbow joint. The velocity at each moment of the swing is, of course, shown by the width of the vibrations on the smoked drum. (In Figure II are shown the curves obtained for one subject with three different weights.) The greater the tension of these muscles, and the less the relaxation, the smaller will be the velocity obtained under a given torque, and the smaller the total angle through which the arm travels. McKinley and Berkwitz supply the physical equations connecting the weights applied, the dimensions of the apparatus, the friction or other errors, and the actual tonus which is to be measured.

In measuring their records, these authors determined the angle at which the greatest velocity was attained. This involves the assumption that the angular velocity of the swinging arm increases regularly to, and decreases regularly from, the maximum. Actually, however, this is the case only when the muscles are genuinely relaxed. The majority of the subjects are unable to avoid innervating these muscles at various points in the swing, and the result is an irregular velocity with several maxima, or a long period of practically constant velocity. Such irregularities are clearly visible in the top two curves of Figure II; with the heavier weights, the increase and decrease of velocities will be seen to be more regular. Thus it seemed safer to work from the actual maximum velocity attained which is, of course, just as directly proportional to the arm muscle tension as is the ideal angle at which this velocity is attained. The error of measurement seems to be smaller, and it is even possible that the higher reliability found by the present writers than by McKinley and Berkwitz, may be due to these diminished errors of measurement.

CHAPTER IV

THE EXPERIMENTS

This chapter contains a description of all the experiments employed, a brief account of relevant antecedent studies, and a statement of the nature, range, and reliability of the measures obtained. The data presented here are the basis for our argument concerning the consistency of habits of gesture; but since no conclusions are drawn in this chapter, the reader whose interest lies in results rather than method may prefer to omit it.

Reading and Counting. Normal speed of saying the alphabet, counting, or reading set passages has been measured by Wagoner and Downey (*191*), Baxter (*8*), Hübel (*73*), and Braun (*20*). In general they found high reliabilities and good agreement with other measures of natural speed of movement.

The Experiments

1. *Reading aloud* (sessions I and III).
   Measures obtained: speed; qualitative observations and ratings on voice intensity, fluency, and movement during speech.
2. *Counting aloud* (sessions I and III).
   Measure obtained: speed.

Materials and instructions. At session I, the subjects were required to read aloud a short story in their natural manner; the story contained 327 words. At session III they read a passage from an elementary text in psychology, 189 words in length. Later in the same sessions the subjects were asked to count aloud in their normal way from 1 to 30 (I), and from 31 to 60 (III).

Range of times and medians.

Reading I: 82.5 to 154.5 seconds; median 105 seconds.
" III: 60.7 to 121.0 " " 70 "

Range of times and medians, continued.
Counting I: 8.2 to 25.9 seconds; medians 16.7 seconds.
" III: 15.8 to 27.5 " " 20.5 "
These times are the averages of the records kept by two experimenters.

Reliabilities and intercorrelations.
Reading: I, III = + .78. Corrected reliability = + .88.
Counting: I, III = + .74. Corrected reliability = + .85.
The correlation between the paired speeds of reading and counting = + .21.

Walking and Strolling. No investigation has been found in which ordinary speed of walking out of doors has been measured. Anders (*5*) and Braun (*20*) have taken records under laboratory conditions, making their subjects walk along a corridor. Anders was chiefly interested in the relation of walking rhythm to pulse and respiration rates. Braun found a correlation of +.72 between the speeds when subjects were directed to walk "as in the street," and when they were told to walk "as in a room." In view of the low correlation between walking and strolling speeds found in the present study, this figure probably includes a considerable spurious element due to experimental conditions. There apparently has been no psychological work on length of stride.

The Experiments
1. *Walking*, out of doors (I and II), indoors (III).
2. *Strolling*, indoors (II and III).
Measures obtained: speed of walking, speed of strolling; length of walking strides, length of strolling strides.

Experimental conditions and instructions. At the end of sessions I and II, subjects were timed as they left the laboratory, until they reached a certain point some fifty yards away. The number of strides was also counted by both experimenters. The subjects were specifically requested to take no notice of the experimenters who were watching them, but to try to proceed as if they were unobserved.

Halfway through session II, subjects were told to stroll round the outside of a large lecture room, not as if they were walking out of doors, but as if they were walking up and down in their own rooms, meditating. At session III, this was repeated in a library (the total distances being the same for all subjects). After this the subjects repeated the circuit, but were instructed to walk as if out of doors, in their usual manner. The lack of correlation between walking and strolling speeds, given below, indicates that there was no "perseveration," or influence of strolling on walking speeds.

Since length of stride is largely dependent on the subject's height, the total paces recorded in each of the five experiments were each multiplied by the heights,—the simplest method for eliminating this factor. The raw measurements were then ranked and correlated as usual.

Discrepancies between the records of the two experimenters were more pronounced in these experiments than in any other. They do not, however, seem to be large enough to harm seriously the reliabilities of the observations.

Ranges. Walking times I and II: 23.0 to 34.5 seconds, median 28.5 seconds.

Walking times III: 23.8 to 38.5 seconds, median 32.3 seconds.

Strolling times II: 23.5 to 50.1 seconds, median 33.0 seconds.

Strolling times III: 34.0 to 129.1 seconds, median 54.5 seconds.

Number of paces, before being corrected for height:

Walking paces I and II: 50 to 68, median 59.

Walking paces III: 45 to 60½, median 53.

Strolling paces II: 36 to 53, median 43.

Strolling paces III: 51 to 90½, median 67.

Intercorrelations.

Speed of Walking: I, II = +.85; I, III = +.49; II, III = +.46. Average intercorrelation = +.60. Corrected reliability of the three measures combined = +.82.

Speed of Strolling: II, III = +.76. Corrected reliability = +.86.

Length of Walking Strides: I, II = +.64; I, III = +.61; II, III = +.40. Average intercorrelation = +.55. Corrected reliability = +.79.

Length of Strolling Strides: II, III = +.68. Corrected reliability = +.81.

Correlation between the paired measure for strolling speed, and

the combined measure for walking speed = −.07; between length of walking and strolling strides (both corrected for height) = +.02.

Estimation of Known Sizes. As early as 1898, Wolfe (*198*) noted that when children were told to draw rectangles the size of dollar bills, over 90% were too small, whereas the estimations of the sizes of small coins were usually too large. Writing of a group of subjects whose average estimations were exceptionally large, he ascribes the result "to the presence in the class of three or four *large-hearted* fellows who do and see all things in a generous way." It is curious to find in the literature of the 90's such a pointed reference to personality as a determining factor in simple psychological judgments. The constant error in reproduction of lengths has long been recognized by psychophysics, but besides Wolfe no one seems to account for these constant errors in terms of personality.

The Experiments

1. *Estimation of distances between hands.*
   Measure obtained: average degree of overestimation of length.
2. *Drawing of circles the estimated size of 25¢ and 50¢ coins.*
3. *Drawing of rectangles the estimated size of dollar bills.*
   Measures obtained: degree of overestimation of these sizes.

1. *Estimation of distances between hands.*

Instructions. During session III the subject was told to stand up and close his eyes; he was then asked to move his hands apart a distance which he believed to equal two feet. The distance was measured without his seeing the ruler. The experiment was repeated with distances of one foot and eighteen inches, and all three observations were then taken a second time. Scoring is given in terms of the percentage of over- or underestimation, the six records being averaged for each subject. (Correct estimates would yield a score of 100%.)

Range of overestimation of distances between hands.
156% to 60.3%, median 98.7%.

Reliability. This experiment was not repeated, but the intercorrelation of the six readings obtained at the single session may be computed. This gives a corrected reliability coefficient of +.91.

2. *Drawing of circles and rectangles, the estimated sizes of coins and dollar bills.*

Instructions. At the beginning of session III, the subject was given sheets of paper 8″ × 5½″, and told to draw, on separate sheets, circles the size of a 25¢ piece and a 50¢ piece. Then, on a sheet 11″ × 8½″, he drew a rectangle the estimated size of a dollar bill (new style). The three drawings were obtained a second time toward the end of the same session. In measuring these areas, for scoring, and in other experiments that involved the drawing of circles, the figures were often by no means regular; so that the following plan was adopted. The longest diameter was found, and this was multiplied by the length of the diameter perpendicular to it which passed through its middle point.

Ranges of estimation of coins and bills.

25¢	14.2 to	2.78	square cms.,	median	5.67	square	cms.
50¢	20.2 to	5.11	"	"	"	10.65	"	"
$1	110.5 to	34.4	"	"	"	75.1	"	"

The actual areas of the coins [1] are 5.76 and 9.0 square cms., so that the median estimates are fairly correct. The size of a one dollar bill is 103 square cms.; all but two of the 25 subjects underestimated it.

Intercorrelations.

(Since the two sets of experiments were performed at the same session, they will be distinguished by the letters *a* and *b* in the following table.)

25¢ *a*, *b*	= +.59	50¢, *a*, *b*	= +.71
25¢ *a*, 50¢ *a*	= +.84	25¢ *b*, 50¢ *b*	= +.84
25¢ *a*, 50¢ *b*	= +.58	25¢ *b*, 50¢ *a*	= +.63

It should be noted that, although the two sets of experiments were separated by half an hour in order to prevent immediate perseveration, yet these repeat reliabilities may be raised somewhat by being obtained at the same session. The table shows also, that the correlation between the estimation of different coins at the same time, is higher than the correlations for the same coin at different times. Average intercorrelation = +.70. Corrected reliability of overestimation on all four original measures = +.90.

Dollar bill *a*, *b* = +.64. Corrected reliability = +.78.

The correlation between the combined estimation of coins and

[1] *I.e.*, of the squares which would just enclose the circular coins.

bills = +.34, so that it is necessary to treat them as relatively independent variables.

Estimation of Distances and Angles. These experiments derive from Münsterberg's work (*116*). He attempted to pull out and push in a sliding scale whenever, in the normal course of daily life, he felt particularly happy, sad, energetic, weary, jolly, or serious. Energetic and jolly states led, on the average, to consistent overestimations, both of flexor and extensor movements. Weary or serious states produced the opposite effect. In pleasurable states the flexor movements were too small, and the extensor too large, both being reversed during unpleasurable states. Münsterberg, however, made no attempt to eliminate the prejudices which might come from expecting such results. But it is a well-known fact that lines drawn by subjects during pleasurable affective states are usually longer than lines drawn during unpleasurable states.

This mode of investigation was first applied to the study of the permanent make-up of personalities by G. W. Allport (*3*). His apparatus was essentially the same as that used in the present investigation. The subject rested his hand on a horizontal board, on which were ruled a series of numbered parallel lines. Starting from the first line (nearest the body), the subject moved his hand out through a given distance to one of the outer lines, drew his hand back, and then tried to reproduce the movement with closed eyes. The error of his reproduction was recorded without his knowing its extent or its direction. In other experiments the movements and estimates were made from the farthest line in toward the body. Many readings were taken in both directions. Allport found constant tendencies, among normal subjects, to exaggerate the movements in both directions, or to underestimate them. But he was unable to connect

these individual differences with personality traits as he had hoped. Among manic and depressed patients, however, a marked differentiation appeared, the former showing a preponderance of overestimations in both directions, the depressives a majority of underestimations.

The Experiments

1. *Estimation of distances outward from and inward toward the body.* Measures obtained: overestimation of distance from body with hands and with legs; underestimation of distances toward body with hand; average speed of estimation.
2. *Estimation of angles with rotating arm.* Measure obtained: overestimation of angles.

The methods. A surface, some 20″ wide by 22″ long, was marked off with 20 parallel lines, an inch apart, and numbered from 1 upwards. During the experiment, line No. 1 was nearest to, No. 20 farthest from, the subject's body. A modification of Allport's technique was introduced. While resting the tips of his fingers on No. 1, the subject was told a number; he observed how distant was the line corresponding to this number, closed his eyes, and then attempted to slide his hand out, away from the body, to this distance. In this way he estimated previously *seen* distance, not previously *moved* distance. The subject was not at any time allowed to know the direction or magnitude of his errors; he took his hand back to the starting point before he was allowed to open his eyes again for a trial with another number. After a series of 5 outward estimates, he placed his hand on the far edge of the board (line No. 20), and then moved it inward, with his eyes closed, to a line whose number had been given him. Two experimenters recorded the estimates to the nearest tenth of an inch, and the second (who was not concerned with giving the instructions) timed each series of readings unobtrusively. The first experimenter tried, of course, to keep the interval between instructions as constant as possible. Large individual differences were observable in what might be called the cautiousness of making the estimates. Some subjects insisted on jumping the whole distance with a single quick movement of the hand; others slid the hand slowly, hesitated, and often changed the position of the hand before reaching a final decision. The total time occupied by each series of 5 readings expressed these individual differences effectively.

At session I, four sets of readings were taken with the right hand,

two sets outward, two inward; 20 readings in all. Later in the same session the whole process was repeated with the left hand. At session III the experiment was repeated with the right hand only.

A similar method was applied to both feet. The subject stood by a wall or chair for support, and slid his foot out along a scale on the floor, through a previously seen distance. Inward estimates were not attempted, but outward readings were taken at sessions I and III. The scale divisions here were 2″ apart, and No. 12 was the highest number given (*i.e.*, an estimate of two ft.).

Angle estimation. For the estimation of angles, the present writers adapted the apparatus first designed by McKinley and Berkwitz (*121*) for the entirely different purpose of recording muscular tension. A series of estimates was made just before the tonus measurements. A full description of the apparatus is given in Chapter III, pp. 62–65. In the present experiment no use was made of the smoked drum, the system of weights, or the timing device. The subject sat with his lower arm fastened to a horizontal board which was pivoted on a vertical axis. The arm could swing freely about the elbow joint, but when all the muscles were relaxed it tended to assume an equilibrium position, making an angle of 90 degrees with the upper arm. A circular scale, on which angles were marked, rotated with the arm, and a stationary pointer, separate from the board and the scale, indicated the extent to which the arm had rotated away from the initial or rest position. The subject was told to close his eyes during the experiment. The experimenter, standing behind him, then moved the board either outward or inward, extending or flexing the elbow joint, through a definite angle; and immediately brought the arm back to the starting point. The subject let his arm rotate passively during this process, but he was then required to try to reproduce the same movement, "by the feel of it." Extensor and flexor estimates were interspersed; the largest outward angle to be reproduced was 75 degrees from the position of rest, the largest inward angle 30 degrees. Each experiment included five readings in both directions. An experiment was performed with right and left arms at session II, and again with the right arm at session III.

It should be noted that, since the arm rotated in a horizontal plane, it was quite unaffected by gravity. Possibly the extreme estimates may have been aided (according to one subject's introspection) by the pressure of the arm muscles against the straps which held it to the board. Some error was introduced, moreover, by the speed with

which the experimenter conducted the initial rotations. No subject was able to relax the muscles completely, so that however much the experimenter attempted to rotate at a constant angular velocity, he was prevented by the irregularities of contraction and relaxation of the subject's elbow joint muscles against which he had to force the required angle of rotation. This experiment, it will be noted, employs Allport's earlier "kinaesthetic" technique, with, however, passive rather than active participation in the first or "guiding" movement, and without the aid of vision.

Scoring. Allport scored his records either according to the total or average over- or underestimated distance, or according to the number of over- and underestimations. In the present study there were available only five to ten readings in each direction at each experiment, so that total overestimation was an unsatisfactory method, since one large overestimation might outweigh four, or even nine small underestimations. Using simply the *number* of overestimations, only a very small range of variation was obtained. So that, although these two methods of scoring were applied and found to yield good reliabilities, a third method was still more satisfactory. By this method the total overestimated distance (or angle) was multiplied by the number of overestimations, the underestimated distance (or angle) by the number of underestimations, and the latter product was subtracted from the former. Correct estimates were always divided between the over- and underestimations.

Ranges of over- or underestimation of distances from body with hand. In terms of the units described in the previous paragraph, this range is from +120.9 to −148.5, median +17.4.

Overestimation of distances toward body with hands: +137.0 to −123.0, median −1.5.

Overestimation of distances from body with legs (8 right-leg or left-leg readings): +68.8 to −108.8, median −4.0.

Overestimation of extensor angles (in terms of similar units to the above): +425 to −200, median −13.

Overestimation of flexor angles: +175 to −85, median +10.

Total times for twenty estimates with either hand (inward and outward, three experiments): 78 seconds to 225, median 112 seconds.

Intercorrelations.

Overestimation of distance from body with hand: R.HI, III = +.58; R.H.I, L.H.I = +.58; R.H.III, L.H.I = +.46. Average

intercorrelation = +.54. Corrected reliability of all three measures combined = +.78.

Overestimation of distances toward body with hand: R.H.I, III = +.65; R.H.I, L.H.I = +.63; R.H.III, L.H.I = +.60. Average intercorrelation = +.627. Corrected reliability = +.83.

Overestimation of distances from body with feet:

R.F.I, III = +.44;	L.F.I, III = +.44;
R.F.I, L.F.I = +.71;	R.F.III, L.F.III = +.74;
R.F.I, L.F.III = +.46;	R.F.III, L.F.I = +.50.

Average intercorrelation = +.55. Corrected reliability of all four measures combined = +.83.

Overestimation of extensor angles with arms: R.A.II, III = +.46; R.A.II, L.A.II = +.42; R.A.III, L.A.II = +.37. Average intercorrelation = +.417. Corrected reliability = +.68.

Overestimation of flexor angles: R.A.II,III = +.31; R.A.II, L.A.II = +.36; R.A.III, L.A.II = +.46. Average intercorrelation = +.377. Corrected reliability = +.65.

The small number of readings is mainly responsible for the lower reliabilities of the rotating arm estimates. Since the estimates with hands involved double the number of readings, we may predict corrected reliabilities of +.81 for extensor, and +.79 for flexor estimates, had the rotations been based on 30 instead of on 15 trials.

Reliability of speed of estimation: R.H.I,III = +.80; R.H.I, L.H.I = +.91; R.H.III, L.H.I = +.79. Average intercorrelation = +.833. Corrected reliability = +.94.

The relations between the inward and outward estimates differ markedly in the hand and the rotating arm experiments. In the latter, we find, as did Allport, that those who tend to overestimate outward (extensor) angles, also tend to overestimate inward (flexor) angles. The correlation between the two combined measures is +.56. While if all six sets of observations, in either direction with either arm, are compounded, the average intercorrelation = +.37, and the corrected reliability = +.78.

The estimates with the hands, however, tend to show the opposite relation, overestimations of outward distances being associated with underestimations of inward distances. We can derive from the hand estimates a compound variable which represents, not the tendency to overestimate in both directions, but the tendency to "centrifugalize" all the estimates, whatever their direction. In

the following table, the combined rank orders for inward hand estimates have been reversed, so that all the measures are in order from greatest to least centrifugality.

	(1)	(2)	(3)
1. Overestimation from body with hand		+.42	+.54
2. Underestimation toward body with hand	+.42		+.45
3. Overestimation from body with legs	+.54	+.45	

Taking all the original 10 rank orders (3 for each hand and 2 for each leg) and compounding them in the centrifugal-centripetal direction, we obtain an average intercorrelation of +.383, and a corrected reliability of +.86.

The correlation of the angle estimates with these hand and leg estimates is negligible.

Arranging Binet Cubes. Employing the five weighted cubes from Binet tests, the writers hoped to obtain measures from the normal speed of performance, from the subject's spacing of the cubes, and from the neatness or disorderliness of the arrangement determined by the deviations from a straight alignment. It was assumed that no subject would have difficulty in judging the correct order of the weights, so no account was taken of their success. But the subjects were given to understand that this correct arrangement was being tested, so that they were unaware of the three indirect measures in which the experimenters were really interested. The experiment was performed at sessions I and III.

Measures obtained: speed of arranging cubes; extent of cubes; alignment of cubes.

Extent of cubes was measured (not in the subject's presence) in terms of the distance between the outside edges of the outer cubes. Neatness of alignment was determined by the greatest deviation of any one of the middle cubes from a straight line drawn between the two outer cubes. This method proved later to be insufficiently accurate, and several more trials would have been necessary in order to achieve satisfactory reliability.

Ranges.
Range of total times: 13 to 54.2 seconds, median 25.0 seconds.
Range of extents: 11.0″ to 4.38″, median 7.25″.
Range of alignment deviations: 0″ to 1.25″, median 0.22″.

Only one subject arranged the cubes leaving no space between them; the total extent, when they all touch in this way, is 4.38″.

Reliabilities.
Times I, II = +.59; corrected reliability = +.74.
Extent I, III = +.63; corrected reliability = +.77.
Alignment I, III = +.19; corrected reliability = +.32.

No further use was made of the alignment measure.

Estimation of Weights and Handshake. The Smedley dynamometer has found two uses in experimental literature. First it is often employed as a measure of maximum strength of grip. Secondly, as described in the previous chapter (pp. 61–62), both Bills (*12*, *13*) and Duffy (*43*) recorded variations in normal or comfortable grip to indicate alterations in muscular tension. The present writers wished to determine individual differences in the strength of this normal grip, differences which might find expression in such every-day activities as the handshake.

Measures obtained: strength of normal grip; over- or underestimation of weights.

Procedure. Without looking at the scale, the subject was told to grip the dynamometer, not with a maximal pressure, but with what seemed to him a comfortable, slow pressure, as if he were shaking hands with someone. Three readings were taken with each hand at the beginning and again at the end of session I. The experiment was repeated once more at session III.

Range of normal grips: right hand 70.3 to 3.7 lbs., median 24.4
left hand 64.3 to 5.0 lbs., median 25.0

These figures represent the averages of the three readings at any one of the three sittings.

Reliabilities. The intercorrelations between the two experiments at session I (a and b), and the experiment at session III are as follows:

R.H.Ia					
R.H.Ib	+.76				
R.H.III	+.64	+.71			
L.H.Ia	+.84	+.88	+.66		
L.H.Ib	+.71	+.92	+.70	+.90	
L.H.III	+.53	+.68	+.92	+.66	+.73

This table throws interesting light on cross-transference and the effect of the time interval on reliabilities (cf. Chapter V, pp. 98–100).

Average correlation for R,L. hands, at 10 secs. interval .... = .893
" " " R,R. or L,L. " 30 mins. " .... = .830
" " " R,L. " " " " .... = .795
" " " R,R. or L,L. " 4 to 8 wks. " .... = .685
" " " R,L. " " " " " .... = .643

That it was impossible to derive separate measures from the two hands is shown by the following average correlations: R,R. +.703; L,L. +.763; R,L. +.760.

The total average intercorrelation = +.75. Corrected reliability based on all six experiments = +.95.

Estimation of weights. At sessions I and III, a 25 lb. spring balance was hung at about the height of the subject's head, in such a way that only the experimenter could see the scale readings. The subject was directed to pull it down, with his right hand, until the pressure seemed to him to be equal to 8 lbs. Readings were also taken for 4 lbs., 6 lbs., and 12 lbs. The scoring was the average percentage of overestimation on the four readings combined.

Range of overestimation of weights: 380.5% to 87.5%, median 157.8%. Only 3 of the 25 subjects gave average underestimations.

Reliability: I, III = +.50. Corrected reliability = +.67. These low figures were due to the large variations of one or two subjects; probably only very few more sets of observations would have been needed to give the paired measure quite a satisfactory reliability.

The correlation between Overestimation of Weights and Strength of Normal Grip = +.38, a figure which is just over three times its P.E.

Tapping and Compression of Stylus. Both the speed and the pressure aspects of tapping have been widely investigated. The methods and some of the results of this work were considered in Chapter III. Braun's study (*20*) most closely resembles our own. His normal

or "most convenient" speed of tapping with the right hand correlated more highly with seven other measures of speed of normal movement than any other single measure (average +.534). And he found this speed to possess a very high repeat reliability over several months of retests. Downey (*40*) obtained a fair correlation between normal speed of tapping and unspeeded handwriting; but most investigators (*e.g.*, Uhrbrock *184*, Lewitan *103*, Reymert *139*) have used *maximum* tapping speed and have obtained positive correlations with other maximal performances (reaction time, writing, etc.). Reymert, Wells (*194*, *195*), Sherman (*165*), Landis and others (*96*) have investigated the variability of tapping speeds, or the dropping off between the beginning and end of a long period of tapping; Ream (*138*) discusses many aspects of speeded tapping. The University of London school has employed normal speed of tapping as a measure of motor perseveration, cf. Lankes (*97*) and Bernstein (*9*). The implications of all this work will be considered in Chapter V.

The Experiments

1. *Finger tapping* (II and III).
   Measures obtained: normal speed, and pressure of tapping.
2. *Hand tapping* (II).
   Measures obtained: normal speed, pressure of tapping, and pressure exerted by the fingers on the stylus.
3. *Leg tapping* (III).
   Measure obtained: normal speed.
4. *Stylus compression* (II and III).
   Measures obtained: normal speed, pressure exerted by fingers on stylus.

The subject was told, in all the experiments, to tap in a normal way, at his most convenient speed, as if he had to continue for a long time. In general the number of taps during a 20 or 30 second period was counted, or kymographically recorded. During these rhythmic activities, a silent stop watch was used for timing instead

of the easily audible time marker. A screen prevented the subject from watching any of the recording apparatus.

During the forefinger tapping, the thumb and the last three fingers rested flat on the rigid framework. The forefinger, slightly flexed, tapped on the pressure board. Records were taken with the right finger at the beginning of session II, with the left finger toward the end of the same session, and with the right finger again at session III.

Only one set of observations (session II) was obtained for hand tapping. The subject held the stylus as if for writing, and tapped on the board from the wrist at a natural speed.

Foot tapping was recorded for both legs at session III, a different task being interposed between the two experiments. The subject stood against a table, and raised the whole leg from the floor at each tap.

A further "rhythmic" experiment was performed with the stylus. Subjects were told to compress and relax it between the fingers at a regular and comfortable speed. This was done with both hands during session II, and again with the right hand at session III.

Ranges of speeds and point and grip pressures. Finger tapping speeds (calculated from the number of taps in the three experiments): 294 to 24 taps per minute, median 86½ taps a minute.

Hand tapping speeds: 194 to 22, median 68 taps per minute.

Leg tapping: R leg, 58½ to 13, median 27½ taps per min.

" " L " 58 to 11, " 30 " " "

Stylus compression speeds: 201 to 22½, median 51 per minute.

Range of finger tapping point pressures, in terms of kymograph deflections: 3.04 to 0.01 cms., median 0.21 cms. (The actual median pressure was approximately 42 gms. Cf. p. 57.)

Hand tapping point pressures: 2.77 to 0.07 cms., median 0.47 cms.

Grip pressure range: 0.45 to 0.00 cms., median 0.025. Naturally the gram equivalents of such minute pressures could not be determined accurately.

Stylus compression grip pressures: 1.60 to 0.07 cms., median 0.30 cms.

Intercorrelations and reliabilities. Finger tapping speed: R.II, III = +.86; R.II, L.II = +.81; R.III, L.II = +.91.

Right-hand tapping speeds with finger speeds: R.II = +.78; R.II = +.91; R.III = +.86.

These figures show that hand and finger tapping cannot be treated as separate variables. Average intercorrelation of all four measures = +.855. Corrected reliability = +.96.

Foot tapping: R,L. = +.93. Corrected reliability of the paired measure based on both legs = +.96. No doubt this is somewhat too high owing to the short interval between the experiments.

Stylus compression speed: R.H.II, III = +.46; R.H.II, L.H.II = +.58; R.H.III, L.H.II = +.58. Average intercorrelation = +.54. Corrected reliability = +.78.

Though these three varieties of rhythmic speeds (finger and hand tapping, leg tapping, and stylus compression) involve entirely unlike muscle groups, yet they give very high positive associations. The correlation of leg tapping with finger and hand tapping = +.91, and the correlations of these two measures with stylus compression are +.66 and +.70. The average intercorrelation of all 10 original measures = +.693, so that the corrected reliability of this rhythmic speed composite = +.95. (Cf. also pp. 100–105 where speed composites are discussed further.)

Point pressure intercorrelations: finger tapping: R.II, III = +.60; R.II, L.II = +.87; R.III, L.II = +.59.

Hand tapping point pressures with finger tapping: R.II = +.71; L.II = +.79; R.III = +.68. As in the case of speed, these four measures from different muscle groups are best combined, giving an average intercorrelation of +.707 and a corrected reliability of +.91.

Stylus compression grip pressures: R.H.II, III = +.16; R.H.II, L.H.II = +.11; R.H.III, L.H.II = +.53. Average intercorrelation = +.267. Since the correlations are too irregular and low, these measures were compounded with other grip pressure measures before any further use was made of them.

Some of the intercorrelations may be noted here:

Finger tapping speed with pressure (combined measures): −.14.

Hand tapping speed with point pressure (original measures): −.20;

Hand tapping speed with grip pressure (original measures): −.26;

Hand tapping grip with point pressure (original measures): +.52.

Only the last of these is statistically significant, indicating a tendency for the two types of pressure to be positively associated (as Bills found). In all cases there is very slight tendency for the slower subjects to exert greater point and grip pressures.

Drawing Squares, Circles, and Other Figures. A considerable simplification of some of the complex factors

in graphology may be secured through an investigation of the drawing of figures and lines. Oddly enough this rather obvious procedure has hitherto been neglected. The following simple exercise will illustrate the possibilities in this method. When a class of students was directed to draw six circles on pieces of paper of identical size, an extraordinary variety of drawings was obtained, including the following: circles of different sizes, separated from one another and symmetrically arranged; circles of the same size, arranged symmetrically; all the circles overlapping with one another; five small circles in one large one; each circle inside the next largest; and so on. It is not our purpose to study here these interesting qualitative patterns. But the quantitative aspects of size, spacing, and speed of performance are so much more accessible in simplified tasks than in ordinary handwriting, that common and simple drawings were employed in this investigation to provide data on graphic expression.

The Experiments

1. Circles drawn on paper, with right hand (III)
2. Squares drawn on paper, with left hand (III)
3. Circles drawn with crayon, R and L hand (I)
4. Squares drawn on blackboard, right hand (II)
5. Squares drawn in sandbox, R and L foot (II and III)
6. Parallel lines drawn on paper (pressure board) R and L hand (II and III)
7. Length of self-rating checks on rating sheet (III)

Measures obtained:

Speed of drawing (1, 2, 3, 4, 5, 6)
Average area of figures (1, 2, 3, 4, 5)
Total space or extent occupied by figures (1, 2, 3, 4)
Relative proportion of unoccupied space (1, 2, 3, 4)
Fewness or crowdedness of parallel lines (6)
Point and grip pressures in drawing parallel lines (6)
Length of check marks (7)

Experiments (1–5).

The subject was given a writing surface of standard size, and either a pencil or a stylus, black crayon or white chalk, and told to draw a certain series of figures in any way he liked, anywhere on that surface. A few subjects had to be told, in addition, that they were not being tested for speed or accuracy, but none seemed to be aware of the areal and spacing features in which the investigators were interested. The measurement of the figures was, of course, never made in the subject's presence. Almost all the experiments were timed unobtrusively by two experimenters.

In experiments (1) and (2) sheets of paper 11″ × 8½″ were used. Six circles were drawn with the right hand, then three squares were drawn with the left hand, side by side. From corner to corner of each square were added diagonals. Both of these were performed a single time, at session III.

In (3) a roll of paper, 20″ wide was hung on a blackboard. Any length which the subject desired could be unrolled, but the initial exposed surface was 30″ in length. The subject drew six circles with black crayon (R.H.), and later in the same session I he repeated the drawing with the left hand.

(4) On a large blackboard, extending the whole length of the room, the subject drew three squares with diagonals, at session II. The right hand only was used.

(5) A wooden box measuring 3′ × 2′ × 4″ was two-thirds filled with moistened sand, the surface of the sand being smoothed. This sandbox was on the floor, close to a wall or chair to which the subject could hold for support. Round his right toe was fastened a strap, through the middle of which a long screw projected below. With this screw the subject drew one square with diagonals (II). At session III the same experiment was repeated with the left foot.

Ranges and intercorrelations for speeds of drawing.

In the first columns of the following table are given the fastest, the slowest, and median times in seconds for each experiment. Then follow the intercorrelations.

	Range	Median	(1)	(2)	(3R)	(3L)	(4)	(5R)
1. Circles, R. (paper)	5.3 to 36.8	9.7						
2. Squares, L. (paper)	13.8 to 79.5	27.1	+.66					

	Range	Median	(1)	(2)	(3R)	(3L)	(4)	(5R)
3. Circles, R. (crayon)	4.7 to 48.5	8.8	+.52	+.57				
Circles, L. (crayon)	7.0 to 84.3	11.4	+.58	+.80	+.74			
4. Squares, R. (blackboard)	12.5 to 46.0	25.2	+.62	+.67	+.60	+.67		
5. Square, R. (sandbox)	7.5 to 23.5	13.0	+.45	+.50	+.60	+.38	+.76	
Square, L. (sandbox)	5.4 to 27.0	12.2	+.51	+.55	+.47	+.54	+.81	+.61

It would be difficult statistically to separate different speed variables according to the different muscle groups. The average intercorrelation of the seven speed measures = +.60, and the corrected reliability of the composite +.91.

Ranges and intercorrelations for average area of figures.

	Range (*sq. cms.*)	Median	Intercorrelations (1)	(2)	(3R)	(3L)	(4)	(5)
1. Circles, R. (paper)	0.60 to 17.2	3.35						
2. Squares, L. (paper)	1.92 to 24.5	10.55	+.17					
3. Circles, R. (crayon)	12.3 to 607.0	51.0	+.50	−.11				
Circles, L. (crayon)	9.2 to 658.0	68.0	+.60	−.19	+.87			
	(*sq. inches*)							
4. Squares, R. (blackboard)	2.44 to 156.3	29.5	+.37	+.15	+.57	+.40		
5. Square, R. (sandbox)	9.0 to 87.7	33.0	+.33	+.09	+.56	+.63	+.57	
Square, L. (sandbox)	6.25 to 115.5	31.4	+.33	+.40	+.36	+.35	+.56	+.71

The lack of agreement shown only by the left-hand squares (2) is readily explained. This was the most difficult and cramped task the subjects had to do. It was far more difficult, for example, than drawing left-hand figures on a large surface with crayon. Like writing, drawing on paper is almost completely dextral. This measure was discarded.

The average intercorrelation of the remaining six areas is +.515, and the corrected reliability of the composite +.87. We may separate out the three blackboard areas (average correlation = +.617), and the two foot-drawing areas (correlating +.71). The corrected

reliability of each of these combined and paired measures is +.83. They intercorrelate to +.58.

Ranges and intercorrelations for total areas occupied by the figures.

	Range (*sq. inches*)	Median	Intercorrelations (1)	(2)	(3R)	(3L)
1. Circles, R (paper)	1.60 to 28.5	8.2				
2. Squares, L (paper)	1.84 to 14.9	7.0	−.03			
3. Circles, R (crayon)	19.2 to 611.0	87.2	+.50	−.02		
Circles, L (crayon)	13.0 to 909.0	108.0	+.52	+.07	+.86	
4. Squares, R (blackboard)	18.6 to 679.0	122.0	−.03	+.18	+.40	+.25

As before, the squares drawn with the left hand bear little relation to the rest; the other four measures, however, give an average intercorrelation of +.417. The corrected reliability for this composite, which we shall call Total Extent of Figures, is +.74.

Ranges and intercorrelations for proportion of unoccupied space.

The amount of crowding, or the extent to which the squares and circles fill up this total space may be expressed simply as the percentage of space occupied by the figures.

	Range %	Median %	Intercorrelations (1)	(2)	(3R)	(3L)
1. Circles, R. (paper)	6.3 to 82.0	58.1				
2. Squares, L. "	48.3 to 79.0	69.0	+.27			
3. Circles, R. (crayon)	7.8 to 92.0	56.3	+.18	+.16		
" L. "	10.9 to 86.2	67.8	+.56	+.23	+.17	
4. Squares, R. (blackboard)	58.0 to 89.5	72.0	+.42	−.07	+.57	+.35

The correlations are less regular here, but again the left-hand paper squares reduce the total reliability; hence they were omitted, leaving a composite measure which we may call Unoccupied Space in Drawing Figures, whose components intercorrelate to +.375, with a corrected reliability of +.71.

Experiment (6): Drawing parallel lines.

A simple experiment was performed with both hands (II), and later with the right hand (III). On a sheet of paper pinned to the

pressure board had been ruled two vertical parallel lines, ½″ high and 3″ apart. The subject was directed to draw a row of lines in the space between these two, and parallel to them, with the stylus. The spacing and speed were left to him. A typical record is illustrated on p. 56; it shows the point pressure at the top, then the time units, and the grip pressure below. Ranges and medians for pressure are stated here in terms of cms. of deflections on the kymograph records. These measurements represent the median height of the tracings from their base lines.

Ranges (based on all three trials).

Speed of drawing lines: 132 to 21 lines per minute, median 58 per minute.

Number of lines in the 3″ space: 6 to 70, median 30.

Point pressures: 2.27 to 0.06 cms., median 0.47 cms.

Grip pressures: 0.25 to 0.01 cms., median 0.06 cms.

Reliabilities.

The several intercorrelations and reliabilities are given in the following table:

	Correlations				Corrected
	R.H.II, III	R.H.II, L.H.II	R.H.III, L.H.II	Average Inter-r.	Reliability
Speed	+.75	+.89	+.71	+.783	+.91
Fewness of lines	+.86	+.87	+.72	+.817	+.93
Point pressures	+.50	+.71	+.57	+.593	+.82
Grip pressures	+.47	+.15	+.10	+.240	

Since the grip pressure measures were too unreliable in themselves they were later combined with other grip pressure measures taken from other experiments in writing and tapping.

Experiment (7): Length of self-rating checks.

During some earlier work, where a large number of paper and pencil tests were applied to the same subjects, Vernon noted that there were large individual differences, of high reliability, in the neatness, or carelessness of the check marks drawn by different subjects. Some always drew irregular circles round the answers which they were checking, or wrote check marks of great length; others drew small, neat check marks or circles. The former seemed to be those who would usually be called impulsive or erratic in everyday life.

In the present experiment, only one set of such check marks was available, namely those made by the subjects in filling out their own rating sheets. These ten checks (cf. the rating scale p. 93) were measured.

The *range* of average lengths was from 2.09 to 0.311 cms. median 0.598 cms.

Repeat *reliability* could not be determined, but the corrected reliability based on the average intercorrelation of the ten measurements was +.96.

Handwriting: Gross Features. A discussion of experimental graphology, with especial reference to the characterological significance of handwriting, will be found in Chapter IX, and in the concluding chapters are described certain experiments on the total expressive value of handwriting. For the moment, however, we are concerned only with the general quantitative features of writing: speed, size, and pressure. Both Downey (*39*) and Saudek (*148, 149, 150*) attach great importance to these coarse measures. For Saudek, the speed of writing conditions every other feature, while in Downey's Will-Temperament tests, normal speed of handwriting is explicitly intended to indicate natural tempo of movement in general. Experiments with ratings and with other speed tests only partially confirm this supposition. Uhrbrock (*184*), Baxter (*8*), and Braun (*20*) have found high reliabilities, but only moderate correlations for speed of writing with various measures of normal speed. The views of other investigators on the pressure and areal aspects of writing will be dealt with more fully in the next chapter, while methods of recording the pressures have been described in Chapter III.

The Experiments

1. Copying prose paragraph (carbon paper method). (I)
2. Writing sentence and signature on pressure board. (II)
3. Writing *eee*'s on pressure board. (II)

4. Pressure of resting hand. (II)
5. Writing sentence with crayon on large surface. (I)
6. Writing "One, Two, Three . . ." on blackboard. (III)
7. Writing *eee*'s in sand with feet. (I)

At session II, the subject wrote first one or more lines of *eee*'s, on the pressure board, using a pencil instead of the stylus, in order to become accustomed to the "give" of the board. Several experiments on the pressure board then intervened before he was asked to perform some normal writing with the stylus on the same board. The subject wrote the sentence: "The quick brown fox jumps over the lazy dog" (which he already knew by heart), and signed his name. A specimen kymograph record, showing point and grip pressures and time marking is illustrated on p. 56.

Certain writing experiments were performed with other, larger muscle groups. During session I, the subject wrote, "The quick brown fox . . ." with black crayon on a large sheet of white paper hung on a blackboard. This performance was timed unobtrusively, and other measurements were made later. At session III, the subject was told to write with chalk on a large blackboard the words, "One Two Three . . ." up to "Twenty," each number with a capital letter, and as a separate word. This was timed as usual by two experimenters; the area was measured and the writing erased before the next subject entered the room.

During session I the pointer was affixed to the subject's right foot. He then wrote three letter *e*'s, continuously, in the sandbox. Later this was repeated with the left foot. This task was performed too quickly to be timed accurately, but areal measurements were secured.

The reader should note that different muscle groups were employed in six of these experiments. The stylus and the pressure board probably cramped the hand to a slight extent, so that different results might be expected from those obtained in copying the paragraph. And the blackboard writing differed from the crayon sentence in that the former covered a much larger area, and usually involved walking, bending, or other movements of the whole body. The crayon sentence required movements only of the arm and hand.

Ranges and intercorrelations for speeds of writing.

The ranges of times for the several experiments were as follows:

(1) Copying prose paragraphs: 116 to 317 seconds, median 142 seconds.

(2) Pressure board sentence (excluding the signature): 15.7 to 37.0 seconds, median 21.9 seconds.
(3) Writing *eee's:* 252 to 42.6 per minute, median 155.5 per minute.
(5) Crayon sentence: 16.6 to 33.1 seconds, median 24.5 seconds.
(6) Blackboard writing: ("One Two . . ."): 56 to 114 seconds, median 76.9 seconds.

In the following table are given all the speed correlations for writing and drawing experiments with hand, arm, or leg. A remarkable degree of positive association is found between these often quite dissimilar tasks. Only one out of the 105 coefficients is slightly negative; 64 of them are three or more times their probable error. The average intercorrelation is +.41, which would give a total composite with corrected reliability +.91. The most profitable subcomposites that may be extracted are discussed in the next chapter.

TABLE OF SPEED OF WRITING AND DRAWING CORRELATIONS

(All positive unless otherwise marked).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Copying prose paragraph I														
2. Writing sentence III	.77													
3. Writing *eee's* III	.39	.42												
4. Crayon sentence I	.62	.46	.38											
5. Blackboard "One Two" III	.71	.57	.48	.65										
6. R.H. drawing circles III	.30	.18	.14	.28	.47									
7. L.H. squares III	.20	.20	.01	.07	.23	.66								
8. Parallel lines R. H. II	.17	.28	.55	.25	.29	.47	.48							
9. Parallel lines R.H. III	.10	.32	.46	.41	.44	.54	.34	.75						
10. Parallel lines L.H. II	.04	.16	.35	.04	.18	.30	.40	.89	.71					
11. Crayon R.H. circles I	.42	.28	.30	.53	.53	.52	.57	.36	.31	.11				
12. Crayon L.H. circles I	.29	.27	.16	.29	.40	.58	.80	.50	.48	.47	.74			
13. Blackboard, R.H. squares II	.50	.46	.27	.39	.52	.62	.67	.31	.46	.12	.60	.67		
14. R foot square II	.49	.39	.29	.23	.30	.45	.50	.13	.06–.12	.60	.38	.76		
15. L foot square III	.56	.62	.24	.38	.48	.51	.55	.34	.45	.27	.47	.54	.81	.61

Ranges and intercorrelations of areas of writing.

In computing the size of writing, no attempt was made to deal with separate letters, as in standard graphological practice. The

average length of the line was in each case multiplied by the total height from the top of the small letters in the first line to the bottom of the small letters in the last line. The ranges of areas were:

(1) Copying prose paragraph: 482 to 113 sq. cms., median 250 sq. cms.

(2) Pressure board sentence (including signature): 47.4 to 19.2 sq. cms., median 30.8 sq. cms.

(3) Handwritten *eee*'s: average length of each letter = 0.868 to 0.15 cms., median 0.407 cms.

(5) Crayon sentence: 232.5 to 28.5 sq. inches, median 79.1 sq. inches.

(6) Blackboard writing ("One . . . Twenty"): 1950 to 181 sq. inches, median 538 sq. inches.

(7) Footwritten *eee*'s (the two feet averaged): 81.3 to 9.2 sq. inches, median 33.9 sq. inches.

The intercorrelations are given in the following table:

	(1)	(2)	(3)	(5)	(6)
(1) Copied paragraph					
(2) Pressure board sentence	+.45				
(3) Handwritten *eee*'s	+.22	−.03			
(5) Crayon sentence	+.48	+.39	+.16		
(6) Blackboard "One . . . Twenty"	+.31	+.23	+.16	+.73	
(7) Footwritten *eee*'s	+.32	+.19	+.24	+.31	+.21

Note that the crayon and blackboard writing give the highest intercorrelation, with a corrected reliability of +.84. The handwritten *eee*'s give quite poor correlations. It seemed advisable to make a composite variable for *Total Area of Writing* from these five original measures whose average intercorrelation is +.362 and whose corrected reliability = +.74.

Point pressure ranges and intercorrelations.

In measuring the writing pressure records, the same plan was followed as with the tapping records, *i.e.*, the median observation was used. The height of the tracings was measured at every second, according to the time marker, omitting only the gap between the end of the sentence and the beginning of the signature. Since the *eee*'s, however, gave comparatively smooth curves, the average of the pressures at each second mark was adopted. The scoring of the carbon paper pressure method has already been described (p. 53). The ranges were as follows:

(1) Copying prose paragraph, average number of sheets penetrated: 22.0 to 9.0, median 14.2 sheets.
(2) Pressure board sentence (height of kymograph curves): 1.94 to 0.245 cms., median 0.60 cms. (*i.e.*, a pressure of approximately 120 gms.).
(3) Handwritten *eee*'s: 1.98 to 0.04 cms., median 0.243 cms.
(4) Resting hand: 2.80 to 0.14 cms. median 1.04 cms.

To the list of intercorrelations are added the coefficients for the parallel line point pressures, the two right-hand measures being combined.

	(1)	(2)	(3)	(R.H.)
(1) Copying paragraph (carbon)				
(2) Pressure board sentence	+.42			
(3) Handwritten *eee*'s	+.09	+.64		
R.H. parallel lines (II, III)	+.42	+.66	+.47	
L.H. parallel lines (II)	+.56	+.89	+.31	+.71

If all these five measures are summed into a composite for *Pressure of Handwriting*, the average intercorrelation = +.517, the corrected reliability = +.84.

No reliability figures are available for the pressure of the resting hand which was recorded at session II. Its correlations with the other pressure composites are given in the next chapter.

Range and intercorrelations of grip pressures.

In writing the pressure board sentence the range of grip pressure was 0.39 to 0.02 cms. deflection, median 0.12 cms. The correlations between all the measures of grip pressure are given in the following table.

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Sentence II							
(2) Hand tapping II	+.49						
(3) Stylus compression R. II	+.12	+.16					
(4) Stylus compression R. III	+.19	+.12	+.16				
(5) Stylus compression L. II	+.16	+.31	+.11	+.53			
(6) Parallel lines R. II	+.43	+.38	+.00	+.07	+.13		
(7) " " R. III	+.45	+.09	+.03	+.42	+.09	+.47	
(8) " " L. II	−.04	+.08	+.12	−.12	−.00	+.15	+.10

It will be seen that the correlations are irregular. The very poor results of the L.H. parallel lines may well be due to the difficulty in holding the stylus in this position, since it was supported from above with a view to facilitating right-hand writing. Omitting this measure, the average intercorrelation of the other seven measures is +.233, giving a corrected reliability of +.68.

Muscular Tension. The apparatus designed by McKinley and Berkwitz and adapted to our study is described in the previous chapter (pp. 62–65). The results may be briefly stated.

Five or six readings were taken for the extensor muscles of the right arm at session II, and repeated soon afterwards for the left arm. Half a dozen more readings were obtained for the right arm at session III. Three weights were available, and were applied according as the arm was very tense, semi-relaxed, or very relaxed. Usually observations with at least two different weights were made in each experiment. The weights were 1368 gms., 863 gms., and 515 gms. Owing to the small extent of inward swing of which the arm was capable, it was not found possible to take satisfactory readings for the flexor muscles.

Since only rank orders for muscular tension were required, no attempt was made to correct for friction, as in the experiments of McKinley and Berkwitz; nor was the actual degree of tonus computed. In order to make comparable the readings with different applied torques, all the readings were translated, by an empirical method, into terms of the velocities obtained when the medium weight was employed.

Range of maximum velocities (average of 15 to 18 readings per subject): 56 to 233 degrees per second, median 96 degrees per second.

Reliability: R. arm II, III = +.74; R. arm II, L. arm II = +.74; R. arm III, L. arm II = +.81. Average intercorrelation = +.76. Corrected reliability = +.91. If instead, the reliability is determined on the basis of the intercorrelation of all the original readings (cf. p. 40), an average of 16.24 readings per subject, the corrected reliability is +.965.

Ratings. The rating sheet, which is given below in condensed form, is self-explanatory. It was meant to cover several of the aspects of expressive movement which were not accessible to objective experimentation:

Name of Person Rated.....Rated by.....Date.....

RATINGS ON MOVEMENT

Directions to Raters: Place a check against the appropriate rank.

1. Consider *4* the proper rating for a person who is *average* in the characteristic to be rated.
2. Avoid the tendency to cluster the ratings at or above average, unless objective evidence compels you to do so.
3. Do not base your judgment on a single act, but on your observation of the person's movements under normal conditions at various times.
4. Try to secure a complete set of ratings on each subject, but if it is impossible to make a fairly accurate judgment on certain items you may omit them.

Then follow the ten characteristics listed below with a scale of seven intervals under each.

1. Voice intensity (loudness)
2. Speech fluency (connectedness, not richness of vocabulary)
3. Amount of movement (gesticulation, etc.) during natural speech or conversation
4. Amount of movement during idleness
5. Freedom of movement
6. Smoothness of movement, rhythm
7. Speed of normal movement, natural tempo of activity
8. Neatness of appearance (total impression gained from all attributes, *e.g.*, cleanliness, clothes, hair, shoes)
9. Consistency of the neatness just rated. Is he *equally* neat or slovenly in all the aspects above mentioned? (Rate on a four point scale)
10. Forcefulness of personality. Would the person *popularly* be called forceful, vigorous, dynamic, vital, striking? In common discourse would he be said to have a great deal of "personality"?

The scale was filled out by the two experimenters after the first session, by two or three experimenters at the third session, and by the subject himself. The subject also obtained, if possible, three outside raters, friends, or associates, who rated him independently. An average of 7.84 ratings per subject was in this way obtained.

Reliabilities were determined by the $\frac{\text{S.D. average}}{\text{S.D. infinity}}$ method (cf. p. 40), with the following results:

1. Voice intensity +.765
2. Speech fluency +.744
3. Amount of movement in speech +.840
4. Amount of movement during idleness +.731
5. Freedom of movement +.677
6. Smoothness of movement +.481
7. Speed of movement +.752
8. Neatness +.852
9. Consistency of neatness +.593
10. Forcefulness of personality +.756

Little further use was made of Nos. 6 and 9, owing to their low reliability. The intercorrelations between the other eight are here given arranged so as to form a rough hierarchy.

	(7)	(5)	(10)	(1)	(3)	(4)	(2)
7. Normal speed							
5. Freedom of movement	+.71						
10. Forcefulness	+.71	+.81					
1. Voice intensity	+.55	+.38	+.50				
3. Movement in speech	+.50	+.43	+.32	+.75			
4. Movement in idleness	+.60	+.42	+.31	+.61	+.65		
2. Speech fluency	+.46	+.62	+.43	+.13	+.21	+.26	
8. Neatness	+.05	+.26	+.09	−.10	+.02	−.05	+.37

The very high tendency of all the categories to intercorrelate positively is disappointing, though not entirely unexpected. It may be due to three main causes. First, there probably is a small positive connection between all these qualities, in the direction of energy, vitality, and activity; this explanation receives some support from the slight tendency of almost all the objective measures and tests to agree more positively than negatively, as will be shown later.

Secondly, there is the familiar "characterial halo," or the good and bad general impression. This seems to have some effect among ratings by friends, for on almost every quality the average rating of all the subjects was above the scale average. But the "characterial halo" probably had less influence on this scale than on most; first, because many of the raters were psychologically trained, and were unacquainted with the subjects except through the experimental sessions; secondly, because the qualities are, for the most

part not particularly desirable, with the possible exceptions of speed, neatness, and forcefulness. Neatness itself has the lowest average positive correlation with the rest.

The final and chief cause of the prevailing positive associations seems to be what might be termed a "psychomotility halo." The raters, no doubt, obtain from the subjects a general impression of motor agility, which they fail to differentiate into the separate categories listed on the rating sheet. This general impression is not based upon personal like or dislike. It will be seen in the next chapter that there are frequent high correlations between the ratings and the tests, but these correlations are not always those which would be expected if the ratings really represented the motor traits which are named in the rating scale. In other words, the ratings do certainly indicate something of the motor nature of the subjects, but the raters have not succeeded in distinguishing sharply the different qualities in the list.

The experimenters had little opportunity for observing the subject's movement in idleness, or his speech fluency. They apparently based their judgments of these qualities upon the observations of which they were more certain, such as speed of movement, voice intensity, and the like. The experimenters were slightly more subject to the "psychomotility halo" than were the other raters or the subjects themselves. The average intercorrelation between all the traits as rated by the experimenters was +.357, as rated by associates, +.324.

The lack of differentiation between the categories shows that it is not possible to use the ratings as validatory criteria for the tests. They will therefore have to be treated as partial indicators in the total assembly of data: at times covering some of the ground not covered by the tests, at other times reënforcing the results of the objective experiments.

Summary. The measurements upon which we are to base our conclusions regarding the consistency of movement are derived in part from the rating scale, but chiefly from our battery of many tests of normal expressive activity. Since the tests as well as the ratings yielded several measurements apiece, and since most of the measurements were repeated or in some other way studied

for their reliability, a large amount of raw material is available and waiting now to be transformed into definite results. Before we attempt this transformation, the following table will provide us with a summary statement of all the correlations at our disposal. These are, of course, all derived from the experiments described in this chapter.

AVERAGE UNCORRECTED COEFFICIENTS

31	original measures correlated with identical measures obtained at different sessions	+.644 ±.083
18	original measures correlated with almost identical or identical measures obtained at the same session	+.750 ±.062
23	original measures correlated with the same measure from muscle groups on the opposite side of the body, same session	+.760 ±.060
20	original measures correlated with the same measures from muscle groups on the opposite side of the body, different sessions	+.621 ±.086
15	sets of average intercorrelations (averaging 3.27 measures in each set), used for forming combined measures	+.624
17	sets of average intercorrelations (averaging 5.89 measures in each set), used for forming composite measures	+.526

CORRECTED COEFFICIENTS

9	ordinary paired measures (stepped up 2 times)	+.810
15	combined measures (stepped up, on the average, 3.27 times)	+.830
17	composite measures (stepped up, on the average, 5.89 times)	+.838
11	combined measures (reliabilities determined not by repetition but by split-half method), including ratings, length of check marks, etc.	+.814
34	variables which comprise especially meaningful unities listed in the *Appendix*	+.813

CHAPTER V

THE EXPERIMENTAL RESULTS: CORRESPONDENCE OF MEASURES

The evidence for our conclusions concerning personal consistency in expressive movement is derived chiefly from the reliability and internal consistency of our measures. At the end of the last chapter a summary statement of all the reliabilities was given; but to bring out the psychological significance of these figures, it will be necessary in this chapter to regroup them, and to assort them in such a way as to reveal the hierarchies which have the greatest self-consistency.

The reader will recall that we have at our disposal *repeat*-reliabilities for practically all the original measures; likewise data on the internal consistency of *combined measures*, which are composed chiefly of performances of the same task with different parts of the body; also data on *composite measures*, which include many similar, but not identical performances with different muscle groups. (Cf. pp. 41–43.) Later in this chapter we shall also construct *group measures* which show an even more inclusive level of consistency. And in addition to these hierarchies, the intercorrelations of *all* variables will be inspected to determine whether there is evidence of all-inclusive consistency, that is to say, of some kind of psychomotor "general factor." *Our argument is that wherever reliability or internal consistency is found, we have a presumption of some kind of harmony or integration in the expressive behavior of the subjects.* In every case the presumption will be examined to determine whether

it is psychologically intelligible as well as statistically sound.

Constancy or Repeat-Reliability. Excepting for a few original measures which were discarded, the *uncorrected repeat-reliabilities* for all experiments were reasonably high. In the case of the discarded measures, the low reliability was due to obvious deficiencies in the experimental procedure, and not to the intrinsic inconstancy of the subjects. The average of the uncorrected repeat-reliabilities (+.684) should be interpreted in the light of the time it took the subjects to perform the experiments which supplied these 49 measures. The average of the median times, exclusive of the instructions and explanations to the subjects, was not more than 30 seconds. There is reason to suppose, therefore, that if more repetitions had been given, spaced throughout several weeks, so that 5 to 10 minutes in all had been spent at any task, and the Spearman-Brown formula then applied, reliabilities would have been secured which would be as high as those of psychometric tests that occupy nearly an hour in time.[1]

We are justified, therefore, in concluding that a subject expresses himself a second time in an identical performance with a considerable degree of uniformity. Our first evidence of consistency, then, is of a specific order, namely, that in a given task our subjects are, to rather striking degree, constant in their performance. *Single habits of gesture, as we have measured them, are stable characteristics of the individuals in our experimental group.*

Temporal Consistency. Since, however, the repeat-reliabilities are not perfect, it is necessary to examine them further to determine more exactly the rôle which

[1] For a discussion of prediction by the Spearman-Brown formula, see pp. 45–47.

the interval between the tests plays. The uncorrected reliabilities of tests given twice at a single session average +.75, of tests given at different sessions, +.644. It appears therefore that conditions operating within the individual in the span of a single session have an appreciable influence. Similarly, when tasks which were performed by the right and left hands (fingers or legs), are compared, correlations for measures obtained up to half an hour apart are +.76, and those obtained a month or two apart average +.621. In general, the shorter the interval between the repeated performances, the higher the reliability.

Although the effects of temporary "set" are thus appreciable, they are not of such magnitude as to reduce the problem of consistency to one of mere mood. Our correlations show that even after an interval of weeks the subjects still tend to occupy the same rank order, and that the motor performances tested show a degree of self-consistency beyond that which is produced by temporary determinants.

Inter-Muscle Consistencies. Next we may note that the average correlations for tasks performed with different muscle groups (+.621 and +.760, according to the shortness of the time interval) are just about as high as the reliability correlations for tasks performed with identical muscle groups (+.644 and +.750). Since these figures are based, in all, on 92 different coefficients, representing records of speed, distance, area, pressure, etc., we have here very strong evidence for inter-muscle integration, not only as a temporary, but as a lasting phenomenon. The finding confirms the evidence of other investigators on cross-transfer, *e.g.*, Bray (*21*), Miles (*110*), Norcross (*123*), and Munn (*117*). Even more analogous to our work is the demonstration by Preyer (*135*) and Saudek (*151*) of the similarity of handwriting

executed with different parts of the body: right hand, left hand, mouth, elbow, or toes.

Not only do the *same* tasks performed with different muscle groups show positive association, but also *different* tasks performed by different muscle groups show marked agreement. If this were not so it would be impossible to construct the composite and grouped measures which are so easily derived from our results. Concretely, this means that there is positive correlation between such tasks as the estimation of angles with the left arm and drawing of check marks in self-rating with the right hand, between the pressure of tapping with the right forefinger and the amount of underestimation of distances using two hands. In fact, the internal consistency of all our more elaborate variables is due to just such remarkable inter-muscle correspondences. These results, therefore, show that the correspondences between our measures are not exclusively a matter of cross-transference of identical tasks, but that they are also products of deeper lying consistencies of the order of "group factors" or psychomotor traits. Since these consistencies are discussed later in detail, it is unnecessary here to do more than to point out the fact that inter-muscle integration as we find it, is something more than the familiar phenomenon of the cross-transference of identical motor habits.

Speed. There are 45 original measures of speed, which have been reduced in the following table to 14 variables: 8 paired measures, 4 combined and 2 composite measures. At the bottom of each column in this table is listed the reliability of the variable, and the number of original measures which have been included in it. The statistically significant positive coefficients have been italicized (*i.e.*, those which are three or more times their P.E.); there are no statistically significant negative coefficients.

INTERCORRELATIONS OF SPEED VARIABLES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Reading														
2. Counting	+.21													
3. Handwriting	+.*55*	+.*45*												
4. Blackboard writing and drawing	+.*41*	+.25	+.*58*											
5. Drawing figures, lines, on paper	+.03	+.11	+.21	+.*56*										
6. Foot drawing	+.*38*	−.02	+.*58*	+.*90*	+.*42*									
7. Finger and hand tapping	−.07	+.01	+.09	+.31	+.*53*	+.*36*								
8. Leg tapping	−.23	+.04	+.02	+.11	+.*40*	+.14	+.*91*							
9. Stylus compression	−.20	+.20	+.02	+.00	+.35	+.08	+.*70*	+.*66*						
10. Walking	−.09	+.34	+.20	−.02	−.11	−.01	+.04	+.07	+.01					
11. Strolling	−.03	+.16	+.19	−.12	−.12	+.13	−.27	−.24	−.08	−.07				
12. Estimating distances with hands	+.20	+.*46*	+.*43*	+.24	+.25	+.15	+.11	+.17	+.*39*	+.23	+.16			
13. Arranging cubes	−.06	−.02	−.16	−.24	−.17	−.11	+.05	+.02	+.33	+.01	+.*38*	+.29		
14. Ratings	−.05	−.06	+.02	−.05	+.06	+.27	+.*40*	+.33	+.*41*	+.*41*	+.22	+.21	+.16	
RELIABILITY	.88	.85	.87	.85	.86	.76	.96	.96	.78	.82	.86	.94	.74	.75
Number of measures combined.	2	2	2	5	5	2	4	2	3	3	3	2	2	7.84

At first sight this table suggests a considerable degree of specificity in the speeds of performance. Upon further examination the reader notes the evidence for two types of consistency. In the first place, each of the 14 variables listed is self-consistent; some of these variables represent paired measures and so express mere constancy (repeat-reliability), but others are combined or composite measures, and their self-consistency indicates that certain rather narrow clusters of speeds or "special" speed factors, clearly exist. In the second place, the italicized coefficients reveal a still broader type of consistency, namely, somewhat general speed factors; these are three in number:

(a) a *verbal speed* composite, made up of (1) reading, (2) counting, (3 and part of 4) handwriting and blackboard writing; *i.e.*, 8 original measures. Average intercorrelation = +.46; corrected internal consistency = +.77.

(b) a *drawing speed* composite, which includes (5) drawing figures, etc., on paper, (6) foot drawing, (rest of 4) blackboard drawing; *i.e.*, 9 original measures. Average intercorrelation = +.58; corrected internal consistency = +.77.

(c) a *rhythmic speed* composite, which includes (7) finger and hand tapping, (8) leg tapping, (9) stylus compression; *i.e.*, 9 original measures. Average intercorrelation = +.76; corrected internal consistency = +.90. The ratings seem to be significantly related only with this composite.

Verbal and drawing speeds are still rather closely connected (+.61), but are almost independent of rhythmic speeds (+.07 and +.19). Perhaps the use of more strictly manual tests (in the place of the drawing) would result in a greater independence of the verbal and manual speeds, and confirm Kelley's (*137* pp. 27 f.) view that "verbal facility" and "manipulation of spacial relations" are unique and independent factors.

Is there evidence for a general trait of speed or a *per-*

sönliches Tempo? The writers have recently reviewed several investigations in which the average intercorrelation of numerous speed measures was around +.45 (*4*); and have pointed out that such figures are sometimes interpreted as evidence for generality of quickness, and sometimes as evidence for specificity, depending on the inclination of the investigator. Additional studies are those of Reymert (*139*) who obtained an average inter-*r* of +.44 between eleven types of simple and choice reaction times, tapping and counting. Between simple reaction times for teeth, lips, fingers, hand, elbow, thumb, and feet, the average *r* was +.61. An ingenious apparatus has been devised by Campbell (*27*) for measuring choice reaction times with any of the major muscle groups. He obtained repeat-reliabilities averaging +.94, and intercorrelations between the legs, arms and fingers of +.785. Uhrbrock (*184*) examined a number of speeded performances (tapping, reading, peg board, etc.), in relation to the Downey speed of writing tests, and found that most of them correlated +.50 and +.65 with a composite of 12 such performances. In Lewitan's (*103*) investigation, good repeat-reliabilities were obtained, and close agreement between right- and left-hand measures of the speed of such performances as tapping, arm movements, ball sorting, triangle sorting, and certain body movements. He concludes, however, that, "there is no general psychomotor tempo of the individual, but only agreements and affinities between comparable movements of the right hand and left hand, ..." Most of these studies dealt with speeded performances, but Braun's (*20*) closely resembled the present investigation. With seven types of unspeeded (*i.e.*, "natural") measures,—tapping, ergograph, walking, writing, copying, reading, and singing, he obtained an average inter-*r* of +.44.

In contrast with the general trend of these results, the average intercorrelation emerging from the 14 speed variables in our present series is only +.168. The difference would seem to lie in the fact that other investigators have either confined themselves to similar objective situations (*e.g.*, varieties of reaction time), or they have allowed a spurious "set" to enter, which has raised their correlations. Braun, for example, obtained all his material at a single session, and used the same instructions throughout: "Do this at your most convenient speed." In the present investigation, however, as little mention as possible was made of speed. Excepting in a few instances, the natural and normal speed was measured in a wide variety of tasks, unknown to the subjects; and the measures were taken over a period of nearly three months.

Our 14 variables might, of course, be stepped up to give a general speed composite with an internal consistency of +.72. But here is a case where the use of correction, however legitimate statistically, seems to distort and misrepresent the psychological situation. It might also be argued that the fact that no very high negative coefficients appear in the above table, may be a kind of evidence for a certain degree of generality in speed. But in our results the subfactors are more arresting, and are founded on evidence that is more conclusive than the somewhat artificial argument in favor of a general speed factor. The verbal, the drawing, and the rhythmic speed composites are self-consistent and reasonably distinct.

The significance of our findings on speed may be summarized as follows: (1) There is no conclusive evidence of a single, pervasive, "personal tempo" as Meumann (*108* II, pp. 61 ff.), Guttmann (*64*), and others have held. Work

supporting such claims have usually been based on too narrow a range of measures, or on tests with whose purpose the subject was too familiar. (2) On the other hand, it is false to conclude that there is a specific speed for each activity and for each occasion. Particularly striking is our evidence for the repeat-reliability of our measures, showing constancy of speed when a single task is performed at two different times. (3) Our results favor the postulating of three "group factors" of speed: verbal, drawing, and rhythmic. Verbal and drawing activity are not entirely distinct, owing probably to the selection of tests used; but both are independent of rhythmic speed. Although our range of tests was broader than that of most investigators, still an even wider sampling might suggest additional speed factors in personality. Our results indicate that instead of admitting a single "personal tempo" it would be nearer the truth to postulate "speed factors" of only moderate breadth.

Intercorrelations of All the Test Variables: the Problem of a General Psychomotor Factor. In the *Appendix* will be found a table which gives the correlations between all the main variables (combined and composite measures) which were described in the previous chapter. Page references are added by the side of the name of each variable, giving the page on which its derivation is mentioned. For clarity's sake, all plus signs are omitted, minus signs being retained, while statistically significant coefficients, positive and negative, are italicized. Corrected reliabilities or internal consistencies, are given along the diagonal. In addition there will be found three extra rows and columns of figures; these represent the correlation of each measure with the three main "group factors" which were isolated later, and which will be described below.

At first sight the table looks complex and illogical. There are many instances where correlations might be expected, owing to the similarity of the tasks, but are not actually found (a point of considerable importance to the theory of expressive movement offered in Chapter VII). On the other hand there are cases where relatively high correlations occur between apparently unrelated measures. The general trend of the correlations considered as a whole, is rather low. Had there been great spread or overlap between the measures, one would expect more significant coefficients.

If 27 rank orders (the final number of variables employed, ratings excluded), were compiled completely at random, and the 350 intercorrelations were worked out, these could serve as a check on the probability of the correlations actually obtained. (Instead of 351 coefficients, 350 are taken, since one of them, between Area of Blackboard Figures and Total Extent of Figures, is spurious.)

In the first column of the following table are given the number of random coefficients which would fall within each category (up to 1.0 times the P.E., up to 1½ times the P.E., etc.), and in the last column are given the numbers of correlations at each level actually obtained.

CHANCE	OBTAINED
175 would be 1.0 or less times P.E.	118
65 " " between 1.0 and 1.5 times P.E.	84
47 " " " 1.5 and 2.0 " "	51
29 " " " 2.0 and 2.5 " "	28
18 " " " 2.5 and 3.0 " "	25
9 " " " 3.0 and 3.5 " "	14
4 " " " 3.5 and 4.0 " "	7
1 " " " 4.0 and 4.5 " "	10
1 " " " 4.5 and 5.0 " "	0
0 " " greater than 5.0 " "	13
350	350

Thus, of the 69 coefficients actually obtained greater than 2½ times the P.E. (*i.e.*, greater than +.31), only 36 would not be found with a chance set of variables.

One general feature stands out, however, and that is a greater number of positive than negative coefficients. If all the variables (omitting ratings, and the speed correlations whose proper directionality is doubtful) are arranged in order from greatest to least (overestimation, pressure, area, etc.) there are 28 significantly positive and 9 negative coefficients among the objective tests; including ratings the ratio is 63 to 13. The median coefficient of the 24 test variables alone (since the inter-agreement of the ratings is, of course, influenced by the halo) is +.05; the quartiles are at +.22 and −.10; and the average is +.055.

It is difficult to decide whether this minute average inter-*r* could be due to some chance experimental influence. More probably, since all the variables are arranged in such a way as to suggest the same logical directionality, there is present a very small general factor of "vitality" or "power" (cf. p. 45 f). Between the sum of the twenty-four tests and the sum of seven ratings there is a correlation of +.49, showing that the psychomotility "halo" and this logical directionality are certainly related to each other. Though we grant that the psychomotility "halo" is an exaggeration of the facts, yet its positive agreement with the results of the tests strongly suggests that the raters were influenced (too greatly influenced) by a *genuine* psychomotor quality of the order of "strength" or "vitality." [1] The separate ratings most

[1] It is tempting to forget our previous condemnation of attenuation, and correct this +.49, *i.e.*, to correlate an infinite number of equivalent tests with an infinite number of equivalent ratings. If we do so, we discover that the theoretical agreement between the "true subjective psychomotility halo" and the "true objective vitality factor" is +.69.

closely associated with the sum of the 24 variables are Voice Intensity (+.55), Movement during Idleness (+.51), and Movement during Speech (+.45). The other ratings correlate only between +.31 and +.21 with the psychomotor "general factor."

"Group Factor" Analysis. A further study of the table in the *Appendix* shows that the correlations tend to be grouped where the variables are psychologically related, in such a way that one is led to suspect that there are present "group factors" of a fairly general order. Three procedures are possible in disentangling such factors, if they exist. These procedures may be termed the logical, the statistical, and the clinical.

The *logical* method consists in arbitrarily grouping the variables which might, on *a priori* consideration, be functionally related. Although this procedure is indispensable, it is never adequate in itself. Used alone it soon leads into difficulties. For instance, who would suppose *a priori* that the Area of a Dollar Bill would not correlate positively with most of the other tests that involve the drawing of figures, or that Overestimation of Angles would correlate quite highly with many of them?

But as was explained in Chapter III, it is also unjustifiable to attempt a purely *statistical* factor analysis (by Spearman's, Kelley's, or Thurstone's methods). In the following work an eclectic and relatively simple procedure was followed. The "group factors" eventually selected include almost all the statistically significant figures, assembled in such a way as to obtain clusters of variables with the highest possible average intercorrelation and internal consistency. In general each variable which was included in a "group factor" and no variable outside, correlated significantly with this factor; though in a few instances (less than 7%) psychological considerations

made it impossible to follow this principle strictly. It is desirable to repeat the warning here that, since Spearman's methods are not applied, the term "group factor" must not be taken exactly in the sense that he uses it. In the present study the term refers to a tentative, broad, group of positively associated variables which may or may not be discovered by other methods than ours.

Thirdly, *psychological* and *clinical* considerations must not be neglected. The variables and "group factors" must *mean* something to us in terms of the personalities to whom they apply. Although it was not the goal of this investigation to validate the measures obtained from the experiments, yet clinical reference to the more distinctive individual subjects does show that most of the variables selected by logical and empirical criteria are at the same time meaningful. (Cf. the motor case studies, pp. 134–146.) Sometimes likewise this clinical method suggested suitable group factors which later statistical analysis confirmed. Furthermore the *naming* of the factors is based upon clinical considerations.

The procedure was therefore one of progressive logical, statistical, and psychological analysis and synthesis. Variables were added to, or eliminated from, the clusters according both to their resemblance and to their correlation with the other components of the cluster. Finally three statistically reliable and psychologically meaningful factors were achieved. These will be described in the following three sections.

Areal Group Factor. The nine variables which best fit together into this factor are the following; by the side of each is given its correlation with the sum of the other eight components.[1]

[1] Since Area of Blackboard Figures and Total Extent of Figures include identical objective elements, their intercorrelation (which is bracketed in the *Appendix*) was omitted in the computation of the average inter-*r* of all the

Area of Total Writing	.69
Total Extent of Figures	.67
Area of Blackboard Figures	.64
Slowness of Drawing	.52
Area of Foot Squares	.48
Overestimation of Angles	.45
Ratings on Movement during Idleness	.39
Length of Self-Rating Checks	.38
Length of Walking Strides	.37

The average correlation between each component and the sum of the other eight (or seven, cf. footnote), is +.51. The average intercorrelation of all nine is +.333, and the corrected internal consistency +.82. It should be noted that the Slowness of Drawing (reversed Drawing Speed) may be, to some extent, a spurious component, in that larger areas of blackboard figures and foot squares require more time. But the correlations of this variable with the other five indicates that it still bears a true functional relationship with the cluster as a whole.

Interpretatively, this group factor signifies an areal—one might call it an "expansive"—motor tendency. It will be recalled that the traditional school of graphologists claims that size of writing indicates such traits as ambition, pride, imagination, manic as opposed to melancholic, characteristics. Downey (*39*) concludes, from a study of these claims, that, in spite of the difficulties in interpretation of size, "an extremely large and free hand may . . . indicate general freedom of impulse, while an abnormally small hand would lead to suspicion of the presence of inhibitory tendencies. . . ." She therefore uses size of script under distraction as one indicator of "motor impulsion." Saudek, however, distrusts size as a graphological symptom, since it is so easily modifiable

nine components of the group factor. Also in this list, their correlations with the remaining seven variables are given instead of with the remaining eight.

at will. Wolfe (*198*) claims that size of drawings of dollar bills is related to personality. But in our study the lack of any statistical association between this measure and an Areal factor made its inclusion impossible. Length of Walking Strides, Overestimation of Angles, Ratings on Movement during Idleness, and Length of Self-Rating Checks seem to fit excellently with the general psychological concept of expansion.

We may examine the factor further by noting its correlations with the other variables which are not included in it. The highest of the positive coefficients (taken from the outside columns of the table in the *Appendix*), are as follows:

Ratings on Movement during Speech	+.35
Fewness of Parallel Lines	+.33
Writing (Point) Pressure	+.26
Overestimation of Distance from Body with Legs	+.21
Overestimation of Distances between Hands	+.20
Area of Coins	+.20
Lack of Neatness Ratings	+.20
and the rest of the ratings (+.28 to +.20).	

Though none of these is as high as three times its P.E., all seem to support our conception of the nature of the factor. As possible exceptions, one should mention the following measures, which might logically be expected to show some relation to the factor.

Area of Dollar Bills	−.06
Length of Strolling Strides	+.09
Extent of Cubes	−.20
Overestimation of Distance from Body with Hand	−.14

With these few exceptions, the statistical evidence supports the logical and clinical. In general, therefore, our interpretation of the areal group factor as equivalent to motor expansiveness appears to be theoretically and empirically valid.

Centrifugal Group Factor

Overestimation of Distance from Body with Legs	.66
Overestimation of Distance from Body with Hands	.55
Extent of Cubes	.53
Underestimation (reverse of Overestimation) of Weights	.53
Verbal Speed	.34
Underestimation of Distances toward Body with Hands	.33
Ratings on Speech Fluency	.33

The average correlation of each component with the sum of the other six variables = +.47, the average inter-r of all seven components = +.298, and the corrected internal consistency = +.75.

This group factor is based chiefly on the centrifugal-centripetal measures. Extent of cubes is apparently more akin to it than to the Areal factor, while the two verbal components may perhaps be considered as representing "verbal centrifugality." The probable reason for the large agreement of Underestimation of Weights with centrifugality is that pulling down the spring balance involved drawing of the hand in toward the body, a distinctly centripetal movement.

Thus the group factor may be interpreted as a general "outward-tendency," freedom and "extroversion" [1] of expressive movement, the reverse of shut-in, restrained, and pedantic motility. A possible confirmation of this interpretation is suggested by a comparison of the standings of the older subjects with those of the younger student subjects. Only one of the nine oldest adults is above the group median in centrifugality, the average ranks of the two groups being 17 (older subjects) and 10.75 (younger subjects). The number of cases is too small to make it worth while determining the statistical signifi-

[1] Whether this motor tendency actually signifies extroversion of personality is beyond the scope of this inquiry.

cance of the difference, but the centrifugal variables are the only ones in which any marked age differences exist. Teachers of classes of various ages sometimes comment on this phenomenon; they obtain the impression of greater restraint and centripetality, both motor and verbal, in adult than in undergraduate audiences. In Downey's experiment she reports that the younger subjects "exhibit more impulsive and free movement" than the older (*39* p. 103).

The chief correlations with other variables are as follows:

Tapping Pressure	−.34
Finger Pressure on Stylus	−.33
Length of Walking Strides	+.27
Fewness of Parallel Lines	−.25
Overestimation of Angles	−.24

Length of Walking Strides may logically be regarded as centrifugal as well as areal. Possibly the two pressure variables correlate negatively owing to the implication of compression, restraint, and therefore of centripetality in heavy tapping and writing. The method of measuring the estimation of angles (cf. p. 73) may account in part for the negative relation of this test, although it might logically be expected to be associated with the centrifugal factor. Other variables which might logically be expected to correlate positively, but which fail to do so, are Length of Self-Rating Checks (−.16), and Overestimation of Distances between Hands (−.04). It should be noted that the areal and centrifugal tendencies, as thus constituted, are quite independent, the correlation between the two group factors being −.03.

As it stands, Centrifugality appears to be somewhat less satisfactory than the Areal factor. But this may well be due to a relative scarcity of tests in our battery which

bear upon the factor of centrifugality. Enough are available, however, to suggest that some fundamental and unified tendency of psychological importance is at the basis of the intercorrelations which were obtained.

Group Factor of Emphasis

Ratings on Voice Intensity	.71
Fewness of Parallel Lines	.65
Ratings on Movement during Speech	.53
Writing Pressure	.52
Overestimation of Weights	.46
Finger Pressure on Stylus	.45
Tapping Pressure	.42
*Under*estimation of Distances between Hands	.42
Verbal *Slowness*	.38
Ratings on Forcefulness	.38
Overestimation of Angles	.36
Pressure of Resting Hand	.32
Unoccupied Space in Drawing Figures	.31

The average correlation of each component with the sum of the other twelve = +.455, the average inter-r of all thirteen components = +.254, and the internal consistency = +.82.

This group factor is much more heterogeneous than the other two, including a wide variety of intercorrelating variables. Their logical interconnection is, at first sight, rather tenuous, but a kind of common factor of *emphasis* seems to run through them all. The components are drawn largely from measures of pressure, intensity, and forcefulness. Fewness of Parallel Lines suggests parsimony of expression, and strong control. It was easy to see during the experimental situations that those subjects who drew few lines were emphasizing them and exerting greater pressure. Unoccupied Space in Drawing Figures means, of course, that the figures were compressed and occupied a relatively small amount of the total space, just as in

the experiment with parallel lines. *Under*estimation of Distances between the Hands would again appear to express firmness and emphasis; it gave a correlation of +.77 with Unoccupied Space, the highest single figure obtained between any two variables. Overestimation of Angles was found to be closely related to most of the pressure measures, though it is less easy to interpret this result satisfactorily.

The group factor includes one or two components from the Areal and Centrifugal groups, and the common variables are probably responsible for the greater part of the correlations with these factors, namely +.25 and −.22.

Other ratings (Speed, Freedom of Movement, and Movement during Idleness) should also have been included on statistical grounds, but it was decided to exclude them owing to their lack of logical connection with the central concept. This cluster of variables suggests rather strongly the general psychomotility factor which, as was shown above, influenced all the ratings, and was probably present to a slight extent in all the tests. Hence the correlation of the components of the factor with the excluded ratings must be to some degree spurious.

With the remaining variables the correlations of the factor of Emphasis are so small that they can seldom be considered as favorable or unfavorable. It may, however, be noted that Muscular Tension correlates +.17 with the factor as a whole.

An important point in connection with this group factor is that the pressure variables alone do not intercorrelate sufficiently highly to give a statistically consistent aggregate. *Mere physical pressure or tension would seem to be significant only as part of a wider and more psychological tendency to make emphatic movements.* This phenomenon is to be noted also in our speed measures. Whereas

there is a certain tendency for them to cluster together, there is little evidence of a uniform general speed factor. Both speed and pressure measures frequently agree more highly with physically unrelated measures than among themselves. This discovery should help to turn the attention of psychologists away from physical categories and toward more strictly psychological categories of movement.

Previous studies of pressure have been made chiefly in experiments on handwriting. In 1907 Meumann (*108*) noted that masculine handwriting showed a curve of pressure with fairly even distribution, and a tendency for pressure to increase as speed increased; whereas the feminine type of pressure is much more irregular, and tends to decrease with increased speed. A. Gross (*61*) found that depressed patients exerted subnormal pressure, and that curves of pressure are to an extraordinary degree characteristic for any single individual. Hirt (*68*) and Diehl (*35*) support these findings, and the latter believes pressure in handwriting offers an index of psychic tension in general. Both Saudek (*150, 152*) and Osborn (*126*) consider rhythmic pressure of handwriting, that is the pattern of variation, to be of fundamental importance, since it is one of the hardest features to disguise and to imitate. The older graphologists (cf. Downey's survey, *39*), interpreted total pressure as expressing the writer's "will" and "energy." Bührig (*25*) calls pressure "the third dimensional element" in writing, but considers it to be of value in the study of personality only in relation to other broader factors. In another field, that of reaction times, Delabarre, Logan, and Reed (*33*) conclude that, "the degree of pressure exerted and the range of its variation are characteristic of the individual. Each has his own special tendencies and his limits of varia-

tion . . ." We have already mentioned (Chapter III) the work of Johnson (*79*) and Duffy (*43*, *44*) on the characteristic pressure curves in tapping and reaction experiments, and the conclusions of Bills (*12*) on the generality and psychological significance of muscular tension and pressure. In brief, many investigators besides ourselves have concluded that "psychological pressure" or emphasis is a stable, and generalized factor in the activities of their subjects.

Are the Results Meaningful as Well as Consistent? The three group factors are novel conceptions in the psychology of expression and movement. As T. H. Pear has pointed out, our vocabulary for the description of motor reactions is exceedingly limited, and this is one of the chief reasons why we are so ignorant of this field. It seems to have been generally assumed (by Taylorism, for example) that movements should be classified according to the categories of physics, namely, in the dimension of speed or time, in the dimension of amplitude or space, or in terms of energy and pressure. The results of our investigation tend to show, on the other hand, that none of these categories are psychologically fundamental, that often measurements have to be regrouped to give logically and statistically satisfactory factors or traits. "Emphasis," for example, partakes largely of the physical dimension of energy or pressure, but seems to involve much more. And the spatial dimension of movement breaks down into two almost independent psychological tendencies, the areal and the centrifugal. Similarly, time (speed) is apparently not an independent general factor; it seems rather to split into special speed factors which in turn correlate more highly with non-speed measures than with each other.

It goes without saying that our discovery of composite

variables and group factors depends upon the actual tests employed, and that other investigators working with a different battery might not arrive at identical classifications. Indeed, certain other wider and narrower composites could be proposed from our own data, for the table in the *Appendix* shows how extremely complex is the overlapping among measures. It must, therefore, be admitted that the group factors we have selected are to a certain degree artifacts. Another investigator might not use the same logical, statistical, and clinical criteria which guided our choice. But it is only fair to point out that the investigation was planned without preconception as to the group factors (if any) to which the results might lead. The battery of tests was constructed entirely from the point of view of diversity, comprehensiveness and applicability, not to prove the existence of any preconceived set of general tendencies.

If the tests were still more numerous, our interpretations would, of course, be more adequate. Probably new group factors would be discovered, and some of the relations now obscure would become clear. As our scheme stands, however, most of the variables find a place. The exceptions, though few, should, for the sake of our theory, receive critical examination. First we may consider the measures which do not correlate highly enough to be included in any of the group factors, and then the coefficients which are statistically significant but fail to find a place in our present classification of factors.

Out of the whole assortment of tests and original measures, only two were discarded as being unfit for inclusion in any of the combined or composite measures. These were the drawing of squares on paper with the left hand, and the alignment or irregularity in the arrangement of cubes (cf. pp. 84 and 77). The first of these, we saw, was a task which was artificial owing to the uneven difficulty ex-

perienced by the subjects in using the left hand. The second was too unreliable chiefly because a very small range of variation was obtained, and only two observations were available. Ratings for Smoothness of Movement and for Consistency of Neatness also were rejected as unreliable.

Ratings for Normal Speed and for Freedom of Movement have not been directly included in any of the group factors, but they correlated positively (+.20 to +.48) with all three; and they have been shown to be closely associated with the other ratings through the psychomotility halo. Neatness has few significant relations, probably because none of the tests bore directly on this psychological tendency.

The Rhythmic speed composite did not find a place in the group factors, but it correlated more highly with speed ratings than did either of the other speed variables. The Strength of Normal Grip (Dynamometer) should also be mentioned. The wording of the instructions was unsatisfactory; nevertheless its only significant correlation, +.38, is with Overestimation of Weights, showing some tendency toward meaningful consistency.

Area of Coins and Dollar Bills may be related to other factors, not covered by the battery of tests and ratings. The former, however, shows consistent but small relations with most of the areal variables. Length of Strolling Strides is unsatisfactory, though its correlations of +.18 and +.15 with the Centrifugal and Emphatic grouped measures are meaningful. Here again the instructions and test situations were artificial. As in the case of rhythmic speeds and the dynamometer experiments, a slight change in the wording might have produced very different results, and it was easy to see that different subjects interpreted the directions differently, and found them somewhat difficult to comprehend clearly.

A few of the other speed tests, such as walking, strolling, estimation of distances with the hands to and from the body, etc., have not been carried beyond the paired or combined measure stage. But these were reliable in themselves; their specificity has already been discussed.

Muscular tension, as mentioned in the previous section, has positive agreements with most of the Emphatic variables (+.17 with the grouped measure). It is probable here that we are dealing with a mainly physiological function, on a much lower level, which could hardly be expected to agree highly with any of the more psychological tests. It will be recalled that the subject was completely passive

during the tension experiments; mental processes only entered in his attempts to relax. There is little reason to suppose that the measure obtained, in spite of the high reliability, possesses any more significance for expression than other merely physiological factors.

There are therefore many indications that the less satisfactory measures are due to faults in the instructions and in the experimental situations, or to the fact that they were taken from a too specific, physiological level. The great majority of tests, diverse as they were, show not only reliability but also significant statistical and logical interrelations with other tests. *At least three quarters of the experiments have been unequivocal in their proof of the interconsistency of expressive movements.* For those measures which are less satisfactory there almost always seems to be an explanation in terms of inadequacy of technique, rather than in terms of specificity in the performance of the subjects.

The same conclusion is reached by a study of the statistically significant correlations between the various tests. With only 25 subjects one might expect a number of "freak" correlations, which may be favorable, unfavorable, or meaningless, in the light of the general trend of the results. A single coefficient which is greater than three times its P.E. proves scarcely anything; but if, say, 80 or 90 out of the 100 highest coefficients obtained are meaningful, the conclusion is inescapable.

Take first the 37 correlations of $\pm.36$ or more between the 24 test variables. All but nine of these have easily fitted into our grouped measures. The nine exceptions are as follows:

Strength of Normal Grip with Overestimation of Weights, $+.38$, is entirely logical evidence for consistency, even though not readily absorbed into the group measures as at present constructed. Area of Total Writing with Extent of Cubes, $-.37$, and with Writing Pressure, $+.45$; these coefficients confirm the discovery that writing area is to some extent emphatic. Areas of Blackboard Figures and Over-

estimation of Distances between Hands, +.38: since both variables partake of areal and centrifugal qualities, this correlation is not unexpected. Length of Self-Rating Checks with Overestimation of Distances from Body with Hands, −.38; a natural enough result when we reflect that the former was found to be positively, and the latter, negatively, related to the Emphatic group measure. Extent of Cubes with Overestimation of Angles, −.37, and with Muscular Tension, −.37; both seem meaningless. Also it is difficult to discern logical relations in the correlation between Area of Dollar Bill and Length of Strolling Strides, −.38, or between Length of Walking Strides and Overestimation of Weights, −.37.

If we include now the ratings and speed variables, but omit the correlations of ratings with one another (which have already been accounted for), 34 more high coefficients emerge, of which all but 15 have been used in the grouped measures. Ten of these 15 are indirectly the result of the close interrelation between the ratings (the halo). For example, if the test for Writing Pressure agrees highly with the ratings on Forcefulness, it will also of necessity agree with most of the other ratings. It would, therefore, be unfair to take such correlations as disproof of the prevailingly meaningful trend.

Ratings on Neatness correlate −.40 with Underestimation of Distances toward Body with Hands. Since lack of Neatness shows some tendency to correlate with Centrifugality, this result is not unmeaningful. The correlation of +.44 between Neatness and Extent of Cubes is to be regarded in the same way.

The speed measures show the following unaccounted correlations: between Verbal Speed and Area of Coins, −.40; between Drawing Speed and Overestimation of Distances between Hands, −.42; between Rhythmic Speed and Overestimation of Angles, −.55. Since, however, speed measures tend in general to be negatively correlated with both areal and emphatic variables, this trend needs no further apology.

We find, then, that most of the statistically significant correlations are likewise meaningful. That is to say, they nearly all indicate unambiguously interrelations between motor performances which cannot be accounted for by chance, nor in terms of the specificity of habit; but which

obviously evince the tendency of expressive movement in one sphere of activity to correspond to expressive movement in other spheres of activity. Of the 71 coefficients which are greater than $\pm .35$ (excluding the intercorrelations of ratings with one another), 47, or 66%, are included in our three group factors. For 18, or 25%, there seems to be a clear explanation in terms of psychological relationship, even though the group factors as they stand do not include them. Only 5 (7%) of our correlations which are statistically "significant" are not at the same time psychologically meaningful.

CHAPTER VI

EXPERIMENTAL RESULTS: PERSONAL VARIABILITY AND CONSISTENCY

Psychomotor consistency has thus far been studied from the point of view of the correspondence of measures in a *group* of subjects. The first two sections of the present chapter will deal with the correspondence of measures within the *individual* subject. The final section of the chapter will show that correlations and statistical measures do not exhaust all our evidence concerning personal consistency, and that very important additional evidence comes from viewing the motor performances of individual subjects in the light of their total personalities, that is to say, as an expression of an individual *style*.

The Statistical Approach and Its Limitations. A simple hypothetical example will serve to explain the statistical measurement of individual consistency. In the table on page 124 are given the raw scores of five subjects (*a*, *b*, *c*, *d*, *e*) on three tests (I, II, III). The same scores are also expressed in the form of rank orders.

Applying the usual methods, we find that the average intercorrelation of the three tests in the whole group is +.40, the corrected consistency +.67. But if we look at the individual scores and ranks, we see that subject *c* obtains the same score in each test, while certain subjects, especially *d* and *b*, vary greatly. For any given level of consistency in the group as a whole there may be wide differences in the consistency or variability of the individuals composing the group. The mean variation (or better, the standard deviation) of each subject's scores

Subject	Scores				Ranks				Mean Deviations of Scores	A.D./M.
	I	II	III	Average	I	II	III	Total		
a	80	85	90	85	2	1	1	1	3.3	.039
b	90	40	65	65	1	4	2	2	16.7	.257
c	50	50	50	50	3	3	3	3	0.0	.000
d	20	70	15	35	5	2	5	4	23.3	.667
e	30	25	20	25	4	5	4	5	3.3	.133

about his own mean may be used as a measure of his variability. But in order to make these indices for different subjects comparable they should be divided by the subjects' means, as in the last column of the table. This procedure gives the Pearson "coefficient of variation." A subject whose simple reaction time, let us say, is 100 thousandths of a second would obviously have much less chance for variation than another with a mean reaction time of 200 thousandths of a second. Several investigations of variability have been weakened by neglecting this precaution.

Thus we see that the calculation of variability for *raw* scores on tests such as reaction time is quite simple. But in many of the tests used in the present study, no absolute physical units are available, and it is therefore impossible to measure individual variability directly; the scores can be obtained only as ranks, percentiles, or in terms of sigmas. Since the scores for a given subject are not all comparable we are forced to deal with his relative standings on the various tests referred to group norms. The problem is similar to the one encountered in the Character Education Inquiry (*66*) where 21 tests of honesty

were applied to over 500 children. The investigators found the best measure of variability for an individual subject to be the S.D. of his 21 sigma scores. This measure was called his M.I. or "mean integration index." In the present study, when physical raw scores could not be used, our variability measures were calculated from the S.D. of each subject's several rank positions.[1] It should be pointed out, however, that a large, unselected population is desirable when using rank or sigma scores for this purpose. Norms that are based on 25 subjects are not likely to be sufficiently representative. This factor therefore introduces a certain error into the results obtained by the method just described.

There have been several previous studies of the variability of an individual's performances. Most of them are concerned with the relation between inconsistency and emotional instability. The attributes of "emotional instability," however, are very miscellaneous, and include such ill-assorted characteristics as normal distractibility, simple alternations of mood, and pronounced psychoneurosis. Actually there is as yet little evidence that measures of variability are diagnostic of emotional instability.

The mean variation of simple reaction times is often claimed to be abnormally high in psychopathological cases. Saunders and Isaacs (*156*), however, who list some thirty references dealing with reaction times in the psychoses, conclude from their own experiments that most normal subjects show practically as great variability. Perhaps a

[1] The Pearson "coefficient of variation" cannot be applied directly to ranks, for a subject who averages 13th out of 25 on several measures obviously has a much greater chance of variation than a subject who averages, say, 5th or 21st. The procedure employed was first to calculate the S.D. of the individual's several ranks about his own mean rank. If his mean rank lay between 1 and 13, then his S.D. was divided by that figure. If it lay between 14 and 25, then 14 was regarded as equivalent to a rank of 12, 15 as 11, and so on. These conversion figures were then used as divisors.

more promising technique, combining the principles of tests for reaction time and cancellation, is that of Kehr (*81*). His curves show greater irregularity of performance among shell-shock cases during the war, than among normal subjects. Wells (*194, 195*) introduced the so-called "fatigue index" as a measure of variability in tapping speeds. According to the work of Landis and his collaborators (*96*) this measure shows some correspondence with criteria of emotionality. The dot-tapping test is also believed by Sherman (*165*) to indicate certain unstable qualities in the most variable subjects. We have already mentioned the findings of Johnson (*79*) and Duffy (*43*) with respect to irregularity of pressure curves in tapping (pp. 59 f., 62).

At higher levels of behavior we have the statement of Saudek (*148, 150*), Downey (*39*), Burt (*26*), and others, that large variations in the size and regularity of handwriting, or fluctuations in slant and alignment indicate unbalanced and dissociated personalities. And it is commonly stated that unevenness in the abilities of a child contributes to emotional instability. The child whose capacities are all at about the same level (whether high or low) is assumed to be the most stable. Spearman (*168*) strongly favors the admission of a general factor of "oscillation" in all cognitive processes, and identifies extreme oscillations with psychopathological trends. He quotes experiments of Flugel which indicate a consistent oscillation factor in various types of cancellation tests, independent of *g* and of perseveration.

There are few investigations which attempt to measure integration (or variability) as a trait of personality. Garrett (*57*) considered integration as the degree of correspondence among various habit-systems in the individual. His material consisted of self-ratings; and the

consistency of scores derived from these ratings gave fair agreements with certain tests such as the Army Alpha, and scales for neurotic make-up and social intelligence. Hartshorne, May, and Shuttleworth (*66*) conclude from their research that the consistency of "character" among their groups of subjects as a whole is very small, although in some subjects the M.I.'s (integration indices) are significant. They interpret high integration indices as indicative of "good character," and present interesting data concerning the correlation of integration with honesty, intelligence, and other desirable qualities.

Hollingworth (*69*), who used a rather different technique in his study of intra-individual consistency, concluded that, while the variability on tests such as tapping, coördination, substitution, and pulse rate was fairly reliable, yet these oscillations were largely specific to each kind of task. Reymert (*139*) analyzed his data on personal tempo from the point of view of the "consistency of variability." He recorded the time taken by each single reaction, and worked out indices of variability on 12 different performances. His average inter-*r* was about +.22. He concludes that one should posit different variabilities for discrete performances (reaction times) and for continuous performances (such as writing or counting).

The fatal flaw in most of the work on consistency and variability, is the low degree of reliability in the measures, and their slight correlation with other types of variability. It stands to reason that the reliability of variability scores should be considerably lower than the reliability of the original test scores. When Hartshorne, May, and Shuttleworth (*66*) compared the M.I.'s of their subjects on half their honesty tests with the M.I.'s as calculated on the other half, the very low corrected reliability or internal consistency of +.40 was obtained.

This survey of statistical studies of individual consistency and variability leads us to five conclusions, all of which bespeak caution in our own use of the method.

(1) When variability is calculated from raw physical scores (reaction times, tapping, muscular tensions, and the like) the measures obtained seem to be reliable in themselves, but to indicate that consistency is specific to particular tasks or types of movement.

(2) When variability is calculated from more complex tests of intelligence or personality whose initial reliability is not high, then the measures of variability obtained are still less reliable. This low reliability calls into question most of the conclusions derived by this method.

(3) The units employed in much of this work, including percentiles, ranks, or sigma scores, are so indirect that their relevance to the problem in hand is doubtful, and the conclusions based on them are insecure.

(4) We must regard as unconvincing the interpretations of *statistical* inconsistency in terms of "characterial weakness" (Hartshorne, May, and Shuttleworth), or "emotional instability" (Wells, Spearman, and others), or even "lack of integration" (Garrett).

(5) In brief, statistical evidence of consistency and variability based on the agreement or disagreement of a single subject's ranks in a variety of tests, is quite remote from our desideratum; it does not constitute genuine psychological evidence for or against our claim that an individual's habits of expression are essentially well integrated.

Statistical Studies of Individual Variability in the Present Investigation. Even in the face of these strictures upon statistical studies of individual consistency, our data offer a favorable opportunity to put the method to further test. It may be that trends will be disclosed

which, however inconclusive in themselves, can be employed as a supplement to the remainder of our results, and as contributory to our argument.

The following variables, each including a fairly large number of measures, and selected from as wide a field of performance as possible, were chosen to provide material for the study.

1. *Overestimation of Angles.* Since both outward and inward estimates correlated fairly highly, and three sets of readings in both directions had been obtained, the six sets of scores were employed. But the separate scores ranged from +425 to −200 (cf. p. 74), so that in this case it was impossible to divide the individual variations by their means. The simple A.D. was therefore substituted as the measure of consistency-variability.

2. *Point Pressures in the Drawing of Parallel Lines.* In the three sessions at least 20 lines had been drawn by each subject, usually more, and their pressures were measured with reasonable accuracy by the kymograph. The usual coefficient of variation was applied here.

3. *Dynamometer.* Six readings of the pressure of normal grip were available, and these could be directly compared with respect to their variability.

4. *Area of Coins.* The percentage of over- and underestimation in the drawings of the 25¢ and 50¢ pieces were strictly comparable as absolute scores.

5. *Muscular Tension.* About 17 observations had been obtained for each subject during two of the sessions. The average variability from the three means was employed, rather than the variation of all the readings from the same mean.

6. *Ratings.* When the 8 most highly intercorrelating items on the rating scale were taken, each individual's variability on his series of averaged ratings (not ranks) could be determined. The measure expresses the degree to which the raters have distinguished between the subject's eight traits.

7. *Writing Speed.* For this, and the remaining variables, rank orders had to be employed. Fourteen orders for speed of drawing and writing intercorrelated sufficiently highly (cf. p. 89) to serve as a source of variability data.

INTERCORRELATIONS OF SUBJECTS' VARIABILITIES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Overestimation of Angles										
2. Point Pressures in Parallel Lines	+.06									
3. Dynamometer (strength of normal grip)	+.21	+.21								
4. Area of Coins	+.29	−.02	−.19							
5. Muscular Tension	−.03	−.10	+.38	−.29						
6. Ratings	+.32	+.21	+.12	+.34	−.37					
7. Writing and Drawing Speeds	+.23	+.13	+.14	+.07	+.03	+.30				
8. Rhythmic Speeds	+.21	−.18	+.28	−.35	+.04	−.02	−.23			
9. Areal grouped measure	−.09	−.44	−.37	+.15	+.15	−.01	+.13	−.07		
10. Centrifugal grouped measure	−.04	−.51	−.26	+.13	−.41	+.26	+.17	+.07	+.53	
11. Emphatic grouped measure	−.11	−.08	+.01	−.11	−.07	+.01	−.05	−.31	+.28	+.40

8. *Rhythmic Speeds.* Six rank orders for speed of tapping (right and left forefinger, right hand, and both legs), and three orders for stylus compression were similarly treated.

9, 10, and 11. The components of the Areal, Centrifugal, and Emphatic group factors were also analyzed for the variability of the individual subjects in respect to the several components of each group.

The intercorrelations between the 11 rank orders for variability are given in the table on page 130.

Clearly there is no evidence here for a *general factor of consistency,* the average intercorrelation being about +.02. But there is a strong suggestion of the presence of two independent composite consistencies. The variabilities of the three grouped measures correlate on the average +.403; while this would give a corrected consistency of only +.67, yet it is rather striking in view of the objective independence of the material from which the separate variabilities were calculated. Nos. 1, 2, 3, 4, 6, and 7 may next be considered; (the reason for omitting Nos. 5 and 8 will be discussed below). Though none of the 15 intercorrelations between these six variabilities are as much as three times their P.E., yet 13 of them are positive, and they average +.161. This would give a composite with the meager internal consistency of +.53. Yet again the fact that there is any relation at all, however small, is promising, in view of all the difficulties enumerated above, and the low reliabilities found by Hollingworth and by the authors of the Character Education Inquiry.

Even more interesting is the fact that the two suggested composite variabilities (the general and specific variabilities, as they will be called) are quite independent, correlating −.14. The data, though inconclusive, seem to indicate that variability may be a consistent trait at a

lower and at a higher level of integration, but not at both levels. The subject who is very variable in respect to one type of specific performance is also likely to be variable in respect to other specific performances, but not in respect to the wider, higher levels of behavior. And the subject who is consistent in his standing on the various components of a general motor trait is likely to be consistent in respect to other such general traits, but he will not necessarily be consistent at the lower, more specific levels.

This interpretation is further confirmed by the fact that not all the components of the general and specific variability composites are independent. The general composite shows small positive correlations with variabilities in Ratings and Speeds of Writing, which are themselves the most general in nature among the components of the specific variability composite. It also shows the largest negative correlations with Point Pressure and Dynamometer variabilities, which are probably the most specific components of the specific variability composite. At a still more differentiated level we find the variabilities of Muscular Tension (5) and Rhythmic Speed (8), which do not agree even with the specific composite. Our findings also fit in excellently with those of Reymert and Hollingworth. The former discovered that the measures of variability in reaction time failed to correlate with the variabilities of more integrated activities such as reading and counting. The latter writer showed that the physiological variability of pulse rate, while highly stable in itself, gave negative correlations with most of his other motor variabilities.

The discussion earlier in this chapter led us to expect that inadequacies in statistical techniques for computing variability would make it impossible for any very definite or convincing results to be obtained. But as far as they

go, the results hang together logically, and in their indirect way supply additional evidence of complex consistencies in expressive movement.

Finally it should be mentioned that there is but little relation between the subjects' scores on these composite measures of variability and their standings on the group factors described in the previous chapter. Centrifugality correlates, however, +.38 with specific variability, and +.10 with general variability; and the Emphatic group factor correlates −.05 with specific variability, and +.26 with general variability.

Psychomotor Congruence: Case Studies. In the previous section we attempted to approach individual consistency and variability by the statistical road. The results, though of considerable interest, were arrived at so indirectly that they seem remote from the active and living consistency which claims our attention in our daily associations with people. When we think of an acquaintance as consistent, we do not ordinarily mean that he would rank, say, at the 45th percentile on all his "specific motor activity," or at the 70th percentile in all "general motor traits." We mean something more subjective but at the same time far less abstract and more convincing, namely that his various reactions constitute essentially integrated and harmonious expressions of his personality.

A rigidly scientific study of consistency, one might claim, should stop when it has exhausted the quantitative aspects of its data. But, as argued in Chapter II, we are required in the interests of adequacy, to supplement our statistical results with a "clinical" picture of certain representative cases. When this is done, it is surprising how much more definite and arresting becomes the evidence for the essential consistency of expressive activity.

These four case studies make use of the ratings, of the

subjects' records on all of the tests, and contain a minimum of subjective characterization, always supported by facts taken from their lives outside the laboratory. These case studies therefore are constructed chiefly of objective evidence. A purely subjective description would, of course, be no advance whatever over the impressionistic and unproved assertions concerning consistency which we rejected at the outset of our investigation. By assembling for each individual subject experimental records of his movement, ratings by his associates, a graphologist's study of his handwriting, and the outstanding facts concerning his intellectual and social behavior, we achieve a psychomotor portrait that is adequate, and to all but an irreconcilable specificist, conclusive. The method of the case study discloses patterns of consistency completely lost to sight in all correlational studies, and demonstrates convincingly that *measures which do not correspond statistically may nevertheless be congruent psychologically.*

Included in each case study is a brief analysis of the subject's psychomotor characteristics as judged from his handwriting. Dr. Saudek rated each subject on a five-point scale for several such characteristics. Since these ratings are intended to be "absolute," not relative to this particular group of subjects, they will be given at the end of each case study without modification. The ratings from handwriting were, of course, made without any acquaintance with the subjects themselves or with their performances on the objective tests used in our experiments. The case studies were written before Dr. Saudek's ratings were received, so that all agreement that appears between them will be in the nature of independent validation.

Subject No. 1. Age, about 45; a vocational counselor who deals only with the placement of business executives.

Gives the impression of being vigorous, brisk but pleasant, efficient and "go-getting." When he commences an interview with a client, he asks, "Where do you want to be five years from now?" He claims to be successful in "sizing up" his clients with the aid of this and a few other questions. He insists that ascendance and extroversion are necessary traits in an executive.

The ratings on this dynamic person place him first in our group of 25 subjects. He represents most highly of all the "vitality" or "power" qualities listed on the rating scale. On the separate traits his standing was: Normal Speed and Speech Fluency (1st place); Forcefulness (1 ½); Freedom of Movement (2); Neatness (3); Movement in Idleness (4½); Movement in Speech (6), and Voice Intensity (6½). In other words he is placed by raters in the top quartile on all eight traits.

On the objective tests his ranks are also high; averaging the 24 variables he ranks 2nd in the group. The experiments in which he scored among the highest 25% of subjects were: Speed of Estimating Distances to and from the Body with Hands, Overestimation of Coins and Dollar Bills, Extent of Cubes, Overestimation of Angles with Rotating Arm, Overestimation of Distance from Body with Legs, Fewness of Parallel Lines, Overestimation of Weights, Tapping Pressure, Pressure of Resting Hand, Grip or Finger Pressure on Stylus (1st place). He is first on the total Emphatic group factor, 7th on the Areal, and 13th on the Centrifugal. Other high ranks (7–10) are: Speed of Walking; Area of Foot Squares and Foot Writing; Length of Walking Strides (in relation to height); and Unoccupied Space in Drawing Figures. On Variability of the specific type he is 24th, *i.e.* the 2nd most consistent in the group.

These measurements are entirely congruent with the

impressions given by the man, with his profession and, in general, with the results of the ratings. Undoubtedly they indicate correctly that his habits of movement are predominately firm, strong, forceful, emphatic, expansive, well spaced, and that his judgment is usually rapid.

The measures which at first sight seem to be somewhat contradictory are the following. His Verbal and Drawing Speeds are very low (23), also his Strolling Speed (24), and Size of Handwriting (25) (cf. the illustration on p. 148), Length of Self-Rating Checks (21), Overestimation of Distances between Hands (22), Length of Strolling Strides (24), and rather high General Variability (7½). But these apparent contradictions become reconciled when we study them in the light of his total personality. Although aggressive and self-confident, the subject has developed caution, precision, and the capacity for delay; when committing himself he must be certain. His records on our tests are congruent with a nature that is careful but decisive when in action. His reading and speaking are slow because his enunciation is exceptionally distinct and precise. The reduction in size of his writing, compared with his more automatic expansiveness of movement, seems to indicate the same quality.[1] The striking difference between his speeds of walking and strolling show this same tendency of modifying any headstrong impulses at will.

There are in short *no* measurements in this case which might definitely contradict a subjective interpretation of this man's personality. They record what common sense indicates, the personality of an assured, incisive, expansive executive, with capacity for prudent delay and caution. The halo in the ratings represents the tendency on

[1] Downey's experiments (*39* p. 111 ff.) show that small handwriting is undoubtedly a valid indicator of attention to detail and a capacity for inhibition.

the part of the raters to make "good errors." Their simplification is in the right direction, and illustrates the sort of economy in judgment known as the "stereotype."

Dr. Saudek's ratings based upon a sample of this subject's handwriting are given below. (In each case a rating of 1 or 2 represents high standing on the first of the two qualities in the pair; 3 is average; 4 and 5 show a definite tendency toward the second of the two qualities.)

Neatness—Carelessness. . . .	2
Consistency—Variability . . .	2
Lightness—Heaviness	2
Centrifugality—Centripetality. .	2
Openness—Reticence	2
Quickness—Slowness	2-3
Boldness—Hesitatingness . . .	3
Dexterity—Awkwardness . . .	3
Expansiveness—Concentratedness	4

The reader will notice that several of these ratings confirm the results of the objective tests and judgments by associates, viz., neatness, consistency (especially Specific Consistency), speed (in which he shows two distinct tendencies), hesitation, and concentratedness. There may be slight disagreement on centrifugality. The other qualities have no direct counterpart among our measures. The impression as a whole is certainly not incongruent, although it might be objected that Saudek's ratings are not as extreme at certain points as the standing on the tests; the direction from the average, however, is generally the same.

Subject No. 2. An artist, 28 years of age, unmistakably Italian in type, though born in America. He is exceedingly animated, talkative, and given to gesticulation; seldom remains seated for long when he talks, but rises abruptly

to have more freedom for emphasizing his speech with gestures. He is short in stature but sturdy in build. His interests are almost exclusively aesthetic; he appreciates beauty openly and vehemently, but is not in the least effeminate. He has a happy and exceedingly sociable disposition.

The ratings on this vivid personality are consistent. He ranks first in the group on Freedom of Movement, on Amount of Movement during Speech and during Idleness; and lowest in the group on Neatness. On Normal Speed and Voice Intensity he ranks 5th, on Forcefulness 6th. His standing on Smoothness of Movement (21) suggests his abruptness. He often makes sudden, impetuous gestures, and shifts the position of his legs and arms. This jerkiness was observed during all the experiments, excepting during tasks that aroused his artistic tendencies. The drawing experiments, for example, were performed with smooth, rounded, rhythmic movements.

The subject ranks 24½ on Speech Fluency. This must be interpreted in the light of an education inferior to the level, abstractness, and quantity of ideas which he tries to express. Since, in boyhood, he spoke an Italian dialect at home, he has an additional verbal handicap. The same comparative immaturity appears in his handwriting (cf. p. 148). Verbal and educational deficiencies, coupled with the animation of his temperament, his sociability and aesthetic interests, probably account for his excessive and forceful gesticulation.

The subject's standing in both the Emphatic and Areal group factors is 4th, a clearly confirmatory result. But on the Centrifugal factor he is 24th. His movements, then, are spacious and vivid, but not directed outward from the body. Although it is not our purpose to correlate these

evidences with "inner" traits we may be permitted to remark that this man, though exceedingly expansive, is not of the administrative type nor in the least egotistic.

He stands in the first quartile in the Area of Writing (blackboard, paper, and sandbox); the size of his pressure board writing (see Figure III) and his overestimation of the Size of Coins were highest in the group. The Total Area or Extent of his figures, the Length of Self-Rating Checks, Length of Walking Strides (relative to height), and his Overestimation of Weights also fall within this quartile. Other scores almost as high are Area of Blackboard Figures, Overestimation of Angles, Fewness of Parallel Lines, Writing Pressure, and Muscular Tension. He is the most variable in the group on the more specific performances, and rather inconsistent (7½) on the measure of General Variability.

The speed tests give interesting results. He is high only in such undirected activities as Rhythmic Speeds (tapping, etc.), and walking. In most other timed activities, verbal, drawing, etc., he is among the slowest. No doubt his language handicap and his unfamiliarity with laboratory tests affect such measures adversely. In applying the tests it was often difficult to keep him to the point. He showed excessive synkinesis while reading or writing, and a tendency to burst into discussion on quite other matters. The fact that he is slow in drawing indicates his obvious aim to secure accuracy of form; and his drawings are perhaps the most accurate in the group. While drawing and painting, in every-day life, he is placid and concentrated. It is apparent that a great amount of "nervous energy" is expended during the less skillful (and to him, less important) activities. Another example of test scores which fail to correspond, but are

none the less congruent, is his standing on Writing Pressure (7), Grip Pressure (11), and Pressure of Resting Hand (19). Being an artist he tends to rest his hand lightly on his work, but to hold and press his brush (stylus) quite firmly.

These results illustrate very well our thesis that there is positive correlation between many measures; and that often where there is no positive correlation there is nevertheless a striking and convincing congruence in the patterns of expression. They illustrate also the fallacy of working with exclusively *physical measures*, such as speed and pressure. This subject, for excellent and obvious reasons, does not show a general speed in all his work nor a general pressure. These purely physical conceptions need to be ordered under much more meaningful and significant *psychological* categories of interests, traits, and abilities.

Unquestionably the dominant characteristic of this personality, in addition to his aesthetic values, is an uninhibited, unself-conscious expansiveness, expressed consistently in speech and gesture. "I like," he says with a spacious sweep of both arms, "to do my paintings on large canvasses." It is difficult to describe fully the "personal idiom" of his creative work, or to prove its relation to his tests and ratings. The reader can judge for himself whether or not the motor qualities just described are congruent with the subject's style of painting, the outstanding characteristics of which are: large canvasses (when he can afford them), rapid execution, much color, predominantly landscape scenes, impressionism rather than line drawing, and much sunlight.

Judging the motor characteristics of the subject solely on his handwriting, Dr. Saudek gives the following portrait:

Openness—Reticence	1–2	
Quickness—Slowness	2	
Expansiveness—Concentratedness	2	
Lightness—Heaviness	2–4	"Variable, irritable"
Centrifugality—Centripetality	2–4	"Normally centrifugal, but centripetal when depressed" "impulsive though self-conscious"
Boldness—Hesitation	2–4	
Consistency—Variability	4–5	
Dexterity—Awkwardness	5	"Sensitive, nervous, moody, mentally refined, but decadent."

These judgments clearly support our findings on awkwardness (jerkiness), lack of neatness, openness, expansiveness, and above all on variability. Most of the verbal characteristics also are applicable. There seems to be disagreement, however, with our personal opinion (though no experiments are germane to this issue), concerning the subject's moodiness. We are inclined to insist that the subject is remarkably buoyant and happy; Dr. Saudek maintains that he is moody and decadent. It should, however, be remarked that Dr. Saudek showed special skill in penetrating the inexperienced script to a discovery of the subject's "mental refinement."

Subject No. 3. A sophomore, aged 18, with a rather contradictory, immature, unintegrated, nature. He is deferential at times, impertinent at others. Intellectually he is very keen and reads widely, but seldom finishes a book; and is doing unsatisfactory college work, owing apparently to his lack of self-discipline in study. He scorns elementary courses and young instructors, thinking that he should be allowed to attack a subject on its advanced levels with the ablest professors at the start. Ambitious to be a doctor of medicine, he still wants to be acquainted with practically every field of knowledge, and buys the new *Encyclopaedia Britannica* as the first step in his college career. He is tall and attractive, and holds his

miscellaneous knowledge ready for purpose of discussion at any time. He is forward in seeking the acquaintance of distinguished professors. This youth might be characterized in a single "reduced term," by saying that he habitually *oversteps*. He attempts too many courses, and has to give up some; jumps into the advanced stages of a subject and has to retreat; pushes himself too far socially, and is judged impertinent.

The only ratings which fall within the highest quartile are Voice Intensity (2), Movement in Idleness and Forcefulness of Personality (both 3). In all the other traits he is rated between 8th and 10th. The scale does not cover all the significant features of this case, but the high ratings are consistent as far as they go. They indicate, for example, the prepossessing character of his behavior, his aggressive speech and his lack of poise or restlessness.

On the Areal group factor he is 1st, on the Centrifugal factor 4th, and 10th on the Emphatic factor. He is first on all the 24 test variables combined. A great many performances indicate unmistakably his unruly expansiveness. He stands in the first quartile on practically all the measurements of area, including: Total Area of Writing, of Blackboard Figures, Foot Squares and Writing, Total Extent of Figures, Overestimation of Coins and Dollar Bills, Length of Self-Rating Checks, also in Length of Walking Strides, Overestimation of Distances from and toward Body with Hands and Legs, Overestimation of Angles, Fewness of Parallel Lines.

The area of his crayon circles was indeed twice that of any other subject. In drawing parallel lines, he started to extend his lines from the bottom to the top of the paper (five times longer than the task required). In writing in the sand, he commenced in the extreme corner, as if to be

able to fill more space. In drawing squares he often prolonged the sides far beyond the corners.

He showed the greatest muscular tension of any subject. Here we have a clear experimental demonstration of the contradiction in his own nature; excessive activity but strain and tenseness, just the combination involved in "overstepping." Also he stands 4th in variability on the more specific performances; and it is not incongruent to find him 20th, *i.e.*, rather highly consistent, on the more general performances. Again, his walking speed is the quickest in the group, but all his other speeds of movement average around 22nd. His medium or low standing on the pressure measurements may indicate lack of confidence in his own extreme movements. In many of the tests he seemed suspicious, regarding them as measurements of accuracy and performing them slowly.

This student has one published literary production, which is the story of a self-conscious and inferior youth who covered up his complete lack of self-possession among the opposite sex, by putting on airs and bragging of fictitious escapades. The story has an authentic ring.

Summing up this case, we may say that there is marked agreement in all the areal and centrifugal measures. But since the pressure measures are on the whole light, his tension great and his judgments slow, the excessive movement must be viewed in the light of lack of confidence and a rather ineffective self-criticism. One who is personally acquainted with the subject has no difficulty in perceiving that this motor picture expresses quite accurately the personality behind it. He is immature, and at the present stage of his development, simply oversteps in the motor sphere, exactly as he does in the intellectual and social spheres.

Quickness—Slowness	1–5	"Inconsistent and unsettled"
Boldness—Hesitancy	1–5	
Centrifugality—Centripetality .	2	
Openness—Reticence	2	
Expansiveness—Concentratedness	2–3	
Lightness—Heaviness	2–3	
Consistency—Variability . . .	4–5	
Neatness—Carelessness . . .	5	"Irritability and restlessness
Dexterity—Awkwardness . . .	5	because of puberty troubles"

There is excellent and almost perfect agreement between these ratings and comments and the results of our study.[1]

Subject No. 4. This case contrasts with the preceding three in that the subject is in no way striking or gifted. A junior in college, he receives low average grades, although he studies fairly hard. He is short in stature, and not distinguished in appearance, manner, or talent. He has an agreeably submissive nature.

In ratings he has no ranks in the first quartile, except on Consistency of Neatness, where he is 4th. Though this measure is unreliable, it is in his case highly suggestive. He stands in the lowest quartile of ratings in respect to Freedom of Movement, Forcefulness (23½), and Speech Fluency (24½). On the three group factors he

[1] One year after this case study was written in precisely the form it is here printed, the subject suffered a severe mental breakdown which at the present time necessitates institutional care. The diagnosis is *catatonic dementia praecox*. In the light of this tragic circumstance the case study acquires new significance. As congruent as the tenseness, extravagance, and irregularity of movement are with the personality described, in retrospect these same motor qualities become actually symptomatic of pre-dementia. The full significance of this erratic movement was not clear to the experimenters at the time the records were taken. The reader, however, will note the shrewdness of Dr. Saudek's judgment that the subject was suffering from "irritability and restlessness because of puberty troubles."

The history of this case also illustrates how the lack of correspondence among our measures in any given case may itself betoken distinctive congruence. The lack of integration in the subject's habits of movement are a reflection of the lack of integration in his personality.

is 20th (Areal), 15th (Centrifugal), and 21st (Emphatic). But perhaps his most striking scores are 21st and 24th on the two composite measures of Variability. All in all he is the most consistent member of the group, according to statistical criteria.

On only two variables in the entire battery of tests does he score in the highest quartile, namely Drawing and Rhythmic Speeds, where he is 5th and 4th. His walking and strolling speeds, and Speed of Estimating Distances, and Arranging Cubes are among the slowest. As for the rest of the measures, they are found for the most part in the third quartile; he ranks 21st on the sum of the 24 test variables. It is the correspondence of nearly all his scores on a fairly low level that is the chief characteristic of this motor picture. In his case there seems to be no especially meaningful patterns among the measures. If an equal degree of correspondence were found in all cases, it would be necessary to expect the existence of a common motor factor, a level of energy or vitality, that conditions all activity. But this simple solution of the problem of expressive movement seems to fit only certain isolated persons, such as the present subject, and perhaps, at the other extreme, Subject No. 1.

We must not fail to note that the three sources of judgment on this personality agree very well: common sense impression, ratings, and objective tests. Whether the methods are controlled or uncontrolled, they agree here uniformly in demonstrating submission and lack of distinction. Using Dr. Saudek's ratings as a fourth criterion we still find essential agreement. He places the subject at a medium or rather low level on all the motor characteristics, and finds the subject more consistent than any of the others.

Neatness—Carelessness	2
Consistency—Variability	2–3
Quickness—Slowness	3
Boldness—Hesitancy	3
Openness—Reticence	3
Lightness—Heaviness	3–4
Centrifugality—Centripetality	3–4
Expansiveness—Concentratedness	4
Dexterity—Awkwardness	4

Matching Records of Movement with Thumb-Nail Sketches. One further experimental approach to consistency led to some striking results. In Figure III are reproduced some of the pressure curves and the handwriting for each of these four subjects. The records are presented in the same order as the case studies. For each subject, (a) is taken from the middle of the experiment in drawing parallel lines. It shows clearly what extreme differences can occur in such a simple task as drawing a straight line one-half inch in length. Part (b) contains the curve of point pressure that was recorded while the subject wrote the three words which are reproduced in (c). (The reader may easily follow the application of pressure for each word, even for each letter. Note, for instance, the crossing of the *t*-bar in "the," recorded as a narrow peak at the end of the first word, especially marked for Subject 3.) There are extraordinary individual differences in the general form, the regularity, and pressure. The reader should notice that the time-marker records below each curve show the speed of writing, but that the horizontal units on the record are different for each subject. The actual thickness of the lines in the tracings is meaningless, since it depends merely on the closeness of the smoked drum to the writing lever. The curves are selected almost at random, as typical pictures of the rhythms of these subjects during spontaneous writing and drawing.

The illustration, as it stands, was shown to a number of judges, students of psychology and adults, all quite inexperienced in graphology, together with the following "thumb-nail sketches" of the personalities.

A. Highly artistic, hyper-active, generous and cheerful, "Bohemian."
B. Colorless, quiet, agreeable, and dependable student.
C. Immature, self-assertive, extravagant, unstable, sophomore.
D. Forceful, active, efficient business man, but cautious and exact.

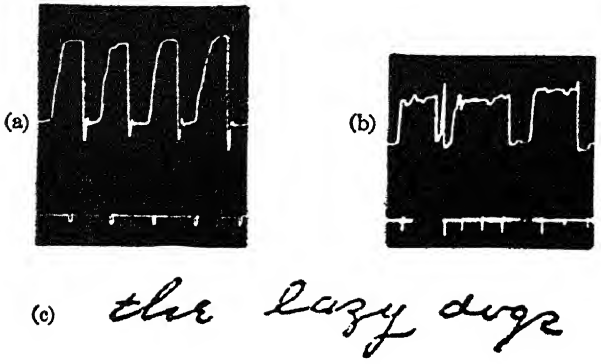
The technique of obtaining the curves was explained to the judges, and they were then asked to match the records with the personality sketches. (The reader will, of course, recognize from the previous case studies that A = 2, B = 4, C = 3, and D = 1.) Can the untrained judge fit these simple records of expression to personalities with whom he is entirely unacquainted except through the half-dozen words of the thumb-nail sketches?

In the following table are given the percentages of matchings of each specimen with each sketch. In all, 32 persons took part, making 128 judgments.

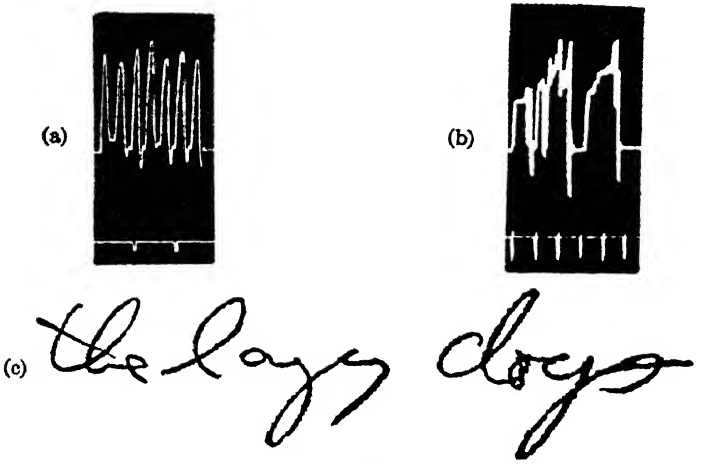
MATCHING THUMB-NAIL SKETCHES WITH RECORDS

		D	A	C	B
Specimens of Curves and Scripts	1	(14.0)	5.5	0.0	5.5
	2	3.1	(5.5)	16.4	0.0
	3	5.5	14.0	(4.7)	0.8
	4	2.4	0.0	3.9	(18.7)

The percentages falling along the diagonal (in parentheses) are correct, totaling 43% of all the judgments. The other 57% are incorrect. Now 25% might be correct

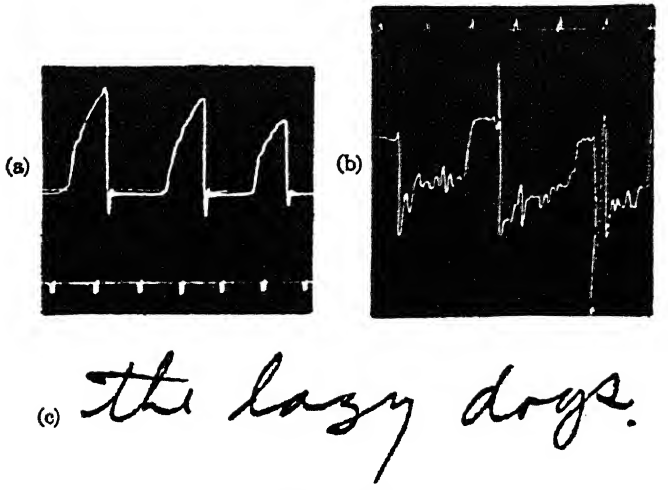


Subject No. 1

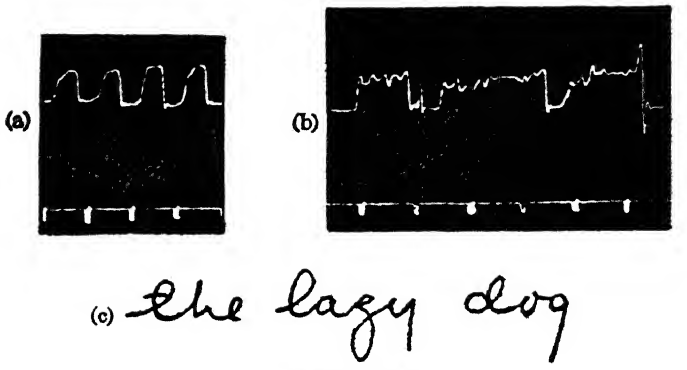


Subject No. 2

FIGURE III. SAMPLE RECORDS OF EXPRESSIVE MOVEMENTS FOR FOUR SUBJECTS.



Subject No. 3



Subject No. 4

FIGURE III (continued).

by chance, but we should remember that, if the judges' guesses had been made at random, we should find, approximately, 6.25% of judgments in each cell. There is a very marked divergence from this random distribution, even if the divergence is not always in the right direction. In other words the judges agree highly in finding the curves and script specimens significantly related to personality. The degree of divergence from random guessing may be expressed satisfactorily by the mean square contingency coefficient. We find that it is +.69 out of a maximum possible +.866.[1] This would be equivalent to a correlation coefficient between +.80 and +.90.

This result proves that the judges agree very closely in regarding the specimens as "expressive." It follows, then, that there must be some logical reason for the differences between this consistent group judgment and the true matchings. The validity (as distinct from the consistency) of the judgments may be found by means of the biserial-*r* method, putting the judgments in the correct cells in one category, and the judgments located in the incorrect cells in the second category. The average percentage of judgments in any one correct cell is 10.75; in any one incorrect cell it is 4.75. The difference in these averages, 6.0%, gives us a biserial-*r* of +.63. The latter figure represents the validity of the matchings.

Now if we examine the incorrect matchings we find that 30.4% of all the judgments consist of confusions of A-2 with C-3. Sixty-one per cent of the judges made this very pardonable mistake. The thumb-nail sketches of the two subjects both suggest expansive and uninhibited natures. Furthermore, the handwriting of A-2 was clearly immature, and "immaturity" is the quality stressed in the

[1] Cf. Yule (*204*).

sketch of C-3. These factors undoubtedly are responsible for the confusion.

No other single error constituted more than 5.5% of the judgments, and most of the mistakes are equally "good errors." Even if the actual number of correct matchings (43%) does not appear very striking (though they represent a biserial-*r* of +.63), yet both incorrect and correct judgments show the tendency of untrained observers to seek (and frequently to find) genuine congruence between expression and "inner" traits.

It is, of course, not to be denied that this group of four subjects is exceptionally heterogeneous from the statistical point of view, and that it is far more difficult to establish these congruences with a more homogeneous group such as might be obtained from a strictly random selection. Because they are somewhat extreme "types," the cases are easier to characterize and easier to match. This consideration, however, does not explain away the conclusions. The expressions and the personalities of other subjects may be less *conspicuous*, but that is no reason for assuming that they are less harmonious. The principle of consistency may hold even in cases where consistency is not apparent to casual observers.

CHAPTER VII

THEORETICAL INTERPRETATIONS

Our experiments have established the fact that there is a considerable amount of consistency in the expressive movement of our twenty-five subjects. This consistency is shown both in the *correspondence* between objective measurements, and in the *congruence* of habits of movement when these habits are viewed in the light of the total personality. It is the purpose of the present chapter to examine various theories of psychosomatic organization to determine with which theory (if any) our results are in accord.

The Theory of Identical Elements. In the language of this familiar theory our results point to a certain spread or "transfer" of the characteristics of movement from one specific act to another. The doctrine holds that such positive correlations as obtain among our measures are due to the recurrence of identical features in the intercorrelating tasks. Let us take the finding that the subjects who overestimate distance from the body with the legs are likely to make the same type of error to the same degree when using the hands. In this particular case, the theory of identical elements seems quite adequate. It holds that the instructions, the *Aufgabe*, or some constant error in the process of judgment, perhaps all of these, serve as identical factors in the two performances, and fully account for the correspondence obtained between the two measures. It is almost second nature for psychologists to accept this explanation. It signifies that there is no process of generalization in mental life, but

that all apparent spread is due to the identical operation of specific habits in different contexts. The following quotation from Stratton (*176* pp. 4 f.) states the theory from the educational point of view:

> The mind is a convenient name for countless special operations or functions. They act almost as though they were insulated from one another; when you have trained one of them, you have trained that limited function and none else. What you do to the mind by means of education knows its place; it never spreads. You train what you train.

From this standpoint our correlations could only be due to the presence of certain identical "limited functions" in any of the series of measurements that correspond. The theory is also clearly expressed in the following quotation from Thorndike (*182* p. 248).

> Training the mind means the development of thousands of particular independent capacities, the formation of countless particular habits, for the working of any mental capacity depends upon the concrete data with which it works. Improvement in any one mental function or activity will improve others only in so far as they possess elements common to it also. The amount of identical elements in different mental functions and the amount of general influence from special training are much less than common opinion supposes.

This theory has been applied to the problem of the organization of personality, with the result that personality is declared to be "a complex of millions of specific habits." The shortcomings of this view have been discussed elsewhere (*4*); but it should be shown in some detail just why the theory is inadequate to account for the consistency we have found in expressive movement.

(*a*). *The amount of "transfer" (degree of positive correlation) is not proportional to the number of identical elements contained in the two series of measures.* If the reader will refer to the *Appendix* and examine the highest intercorrelations he will find that they are obtained

from such dissimilar variables as the spread of cubes arranged on a table with the hands in a direction parallel with the body, and the overestimation of distances outward from the body with the legs ($r = +.59$). To take another illustration, the overestimation of the distance between the hands correlates highly with the crowding of figures in drawing ($r = +.77$). The *Appendix* furnishes many other illustrations of correlations between unlike functions that are as high or higher than the correlations between very similar functions. Conversely, very low correlations are frequently met where the objective identity of the situations would lead one to expect high positive relations. For example, such similar functions as the drawing of a rectangle the size of a dollar bill and circles the size of standard coins correlates only to the extent of $+.34$; overestimation of angles (from the elbow) and overestimation of distances with the hands (out from the body) correlate $-.02$; pressure of resting the hand, and writing pressure agree likewise only to the extent of $-.02$. In short, there is little tendency for functions to correlate in proportion to the number of single features which they have in common. It should also be noted that when absolute identity is aimed at, as in the repetition of precisely the same task for the purpose of determining reliability, the coefficients are far from perfect; many of these are actually less than the coefficients obtained between unlike functions.

(*b*). *The concept of "element" is too ambiguous to have significance in a theory.* How inclusive may the specific function be and still be known as an element? Let us examine for a moment the correlation of $+.59$ between the extent of the span of the cubes arranged with the hands and the overestimation of distances with the legs. Clearly the overlap cannot be due to elements in the

sense of muscle fibrils, neurones, instructions, or apparatus, for none of these are identical in the two performances. The only possible "elements" common to the two are mental sets or attitudes, specifically in this case, a disposition to make centripetal or centrifugal movements, whether with the arms or with the legs. To call this inclusive and generic determinant an "element" is paradoxical.

(*c*). *Similarly, the concept of "identity" is vague and valueless.* How identical must two "elements" be to cause transfer? Absolute identity cannot be expected, for no two stimuli are ever precisely the same in their effects; their objective contexts alter them, as do the infinitely complex and shifting intrinsic conditions of the organism itself. Thorndike inclines to regard identity as in part at least neural. "By identical elements," he writes, "are meant mental processes which have the same cell action in the brain" (*182*, II, p. 359). Lashley's rejection of this view is emphatic. "There is no evidence," he writes by way of summarizing his experiments, "to support this belief in the identity of nervous elements. On the contrary, it is doubtful if the same neurones or synapses are involved even in two similar reactions to the same stimulus" (*98* pp. 172 f.).

If correspondence between measures cannot be explained by literal partial identity, can it not be explained by complex similarities either in the stimulus situations or in their effects upon the subjects? Certainly, but complex similarities are not the same as identical elements.[1] For example, the chickens of Köhler's experiment (*88*) always responded to the *brighter* of two stimuli in various pairs (a kind of complex similarity) and not to a specific

[1] Emile Durkheim's discussion of this point (*47*) is still very much worth reading.

brightness to which they had presumably been conditioned. Responses to *relative* brightness difference obviously cannot be expressed in terms of identity.

Let us take an illustration from our own results. The length of the walking stride and the area of the subject's normal writing correlate, $r = +.46$. In such dissimilar fields of behavior it clearly cannot be a question of identical nerve processes. It seems that each subject is simply maintaining what for him is a suitable and congenial level of activity. Somewhere between the minimum and maximum possible extent, there is an extent of movement natural to the subject, and this remains proportionately constant in very diverse situations whatever musculature is employed. The theory of identical elements is not equipped to handle a case of preserving a congenial proportionality in movement, any more than it is equipped to deal with the chicken's responses to proportional brightness differences.

(*d*). *The theory fails to account for congruence of movement in the individual.* Subject No. 2, it will be remembered, ranks high in the areal or expansive measures, but low in the centrifugal. He tends, for example, to overestimate the distances between his hands but to underestimate the distances away from the body. Objectively considered there are many possible identical elements between these two tasks; but his behavior runs counter to the theory. The fact can be best accounted for (since other measures bear out this single experiment) by assuming that he *is* an expansive but not a centrifugal person in his motor activity, that he has a motor *trait* of expansiveness but not of centrifugality. To take another example, the same man displays random and excessive movement during writing and reading, but is placid and concentrated when drawing. This utterly different be-

havior in the face of situations possessing apparent "identical elements" can be explained only by referring to his lack of interest and training in verbal matters and to his skill and satisfaction in drawing. It is the man's values and not identical elements that are here in question. The subject is an artistic, lively, expressive, energetic, cheerful individual; not however emphatic, aggressive, egotistic, or executive. His performances are accounted for better if they are regarded as manifestations of these general dispositions of his personality than if they are referred to mechanical factors common to the various tests. Neither the concept of "identity" nor of "element," nor the basic intent of the theory, can adequately express a situation that is so clearly freighted with interest, values, organization, and meaning.

Summing up, we may say that the theory fails to account for either phenomenon encountered in our studies: *correspondence* or *congruence.* Psychomotor organization cannot be explained by "insulated" habits functioning in different contexts by virtue of some element that they have in common. The "spread" seems to be due rather to general motor dispositions, related somehow to the total personality, which saturate specific acts with a common quality. But before examining theories which allow for adequate generalization in expression, we should consider one theory which is closely allied to the one we have just rejected.

The Theory of " Unique Traits." Professor Paterson's *Physique and Intellect* illustrates the position of those who think of personality in terms of unique or unitary traits. Paterson first examines in detail the experimental evidence concerning physical correlates of intellect and of temperament, and finds the evidence predominantly negative. He is accordingly led to two conclusions. First,

intellect and temperament (which he roughly identifies with personality) are composed of unitary traits. He holds that there is little consistency within the structure of personality itself. Secondly, these "unique traits" of intellect or personality show no appreciable correspondence with physical expression. He writes, "the facts in the preceding chapters of this book should go a long way toward counteracting the view that there is a functional unity between mind and body" (*128* p. 272).

Paterson's second conclusion is, of course, not acceptable by the canons of common sense, nor is it entirely justified by the standards of critical experimentation. A certain crusading spirit which is variously directed against "faculty psychology," "formal discipline," and "seductive verbalism," makes many writers oppose popular conceptions of the unity of bodily and mental processes, but renders them correspondingly blind to what may be a residual truth. There is something extravagant in the statement that the meager work at hand goes "a long way toward counteracting the view that there is a functional unity between mind and body." Common sense does not easily surrender as gross illusion the conviction that we can somehow tell a melancholy nature by a melancholy posture, or a kindly disposition by a smiling face.

In reading the evidence upon which Paterson bases his conclusions, one is impressed with the crudeness of the conceptions behind most of the research which he reports. Granted these investigations prove that blonds are not aggressive, and that hydrogen-ion concentration in saliva is not associated with good-naturedness, they provide nevertheless only trivial and inadequate grounds for rejecting more fundamental, if more subtle, types of psychosomatic consistency. The available experiments

are too limited in conception and in scope. Usually they compare only one very narrow aspect of physique with another equally narrow, and arbitrarily conceived, aspect of personality. There emerge array after array of negligible or low positive correlations which indicate nothing more than a naïve formulation of the problem. What is forgotten is that segmental studies necessarily lead to segmental findings. Until the problem is formulated so as to allow *complex* physical expressions to show their *complex* relations to the *inclusive* dispositions of personality, an exaggerated belief in the independence of physical expression and personality will remain.

There is a criticism that must be made against Paterson's treatment of these researches concerned with the physical expressions of personality. In examining their results, he attempts to explain away one low positive correlation after another (the studies he cites infrequently yield zero correlations). It would have been quite as legitimate a procedure for him to seek to excuse the lowness of the correlations obtained. If there are spurious factors operating to raise the coefficients above chance, so are there spurious factors which keep the coefficients artificially low. The method of these studies usually involves first the measurement of some narrow physical characteristic (for example, the concentration of creatinine in the urine), and then the correlation of this variable with ratings on some rather ill-defined trait, such as excitability. Low correspondences are not surprising. For one thing the criterion of ratings is notoriously unreliable. Even when tests of personality or intelligence are used for validation, the criterion still leaves much to be desired. This consideration tempts one to manipulate the coefficients (*e.g.*, by correcting for attenuation) to make them higher. It is, however, not considered good form to do so; but it

is no more unjustifiable than to point, as Paterson so frequently does, exclusively to the extrinsic influences which operate to make the coefficients spuriously high.

The theory of "unique traits" holds not only that physical traits are independent of mental, but that traits on either level are independent of each other. We have just criticized as premature Paterson's attack on the "functional unity between mind and body." It remains for us to examine directly the question whether traits on the motor level are unique.

If the theory of "unique traits" applied to our field of study means that expressive movement is an assemblage of "millions of distinct habits," we must of course reject it.[1] Habits are chained, integrated, combined, and patterned. Instead of single, unassociated habits of movement, we find an intricate interweaving. The great majority of correlations between our original measures are positive; and even when these are grouped into broader clusters, there remains between most of these clusters some residual correlation. It is quite apparent that neither the original measures nor the combined, composite, and grouped measures, are entirely unique.

Uniqueness when it appears in experimental data is to a considerable extent an artifact. In treating our own data in Chapter V, for example, we deliberately grouped related measures in such a way as to secure the maximum independence among the composites. This procedure resulted in the *construction* of somewhat unique psy-

[1] Strictly speaking, a "unique" trait may be a very broad trait, such, *e.g.*, as our "Emphatic group factor" (cf. pp. 114 ff.). Even when conceived in this way the notion has a very limited justification as the following paragraph will show. It is a fact, however, that most of the work which claims to support this doctrine leans to the side of extreme specificity. The majority of the experiments reported by Paterson cover such small segments of the field of behavior that the uniqueness implied is a uniqueness of narrowly specific, and not broad, dispositions.

chomotor "traits," but not in their direct discovery. The method is legitimate enough so long as we wish to see what the maximum independence is that *can* be extracted statistically, but it provides no ground whatever for declaring uniqueness to be an ultimate fact of human nature.

The one-sidedness of the statistical approach is clearly seen when it is compared with the case method. The numerous experiments cited by Paterson are based on an analysis of a whole population of subjects. In the preceding chapters we have at some length discussed the insufficiencies of mass investigations. Traits discovered in populations are not necessarily the same thing as traits discovered in clinical studies of single individuals; nor does uniqueness established by one method necessarily forecast uniqueness by the other.

We have stated briefly our objections to Paterson's position regarding the lack of functional unity between personality and its physical expression, and to his theory of unique traits. Paterson's book has been mentioned explicitly because it contains so clear a statement of what is undoubtedly the predominating view among American psychologists. Another influential work pointing in the same direction is the *Studies in the Organization of Character* (Character Education Inquiry, *66*). It too supports the theory of specific, unique traits. Since in certain respects our own problem is similar to that of the C.E.I., and since our conclusions are quite opposed, it will be worth while to suggest two or three differences in approach and outlook which may in part account for the divergent interpretations.

The authors of the *Studies in the Organization of Character* were working with certain aspects of *character*, we with certain aspects of *personality*. The concept of

character is far more difficult to approach, for, as these investigators have shown, it involves considerations of ethical codes, general intelligence, home environment, cultural status, vocabulary, and classroom morale. Such a conception is more sociological than psychological, and lack of "organization" is certainly to be expected in such a mixture of factors; but the specificity discovered should not be interpreted as a specificity of personality.

The subjects in the C.E.I. were children in elementary school. Ordinarily less integration is to be expected in children than in adults, and indeed the results show the degree of integration to be somewhat higher with advancing age (*66* Ch. 22). Our subjects, of course, were all adults.

There is a subjective factor involved in the interpretation of positive correlations in the middle range, and this subjective factor is of great importance in all disputes concerning specificity *vs.* generality. For example, whether or not a coefficient of +.361 between "inhibition" and "honesty" is so meager as to "seem lamentable from the standpoint of defenders of much current theory and method in character education" (*66* p. 152) is a matter of opinion. The issue involved, of course, is the magnitude of a coefficient required to render a correlation "significant." If one demands as a measure of generality correlations approaching unity, a coefficient of +.36 is inconclusive; but if one has in mind all the conditions of research which make it difficult to penetrate the true and intricate nature of personal consistency, such a coefficient may seem to indicate not specificity but a promising degree of generality. Whether it is "high" or "low" depends upon one's point of view. The unreliability of the measures, unnatural homogeneity among the subjects, and the obscuring of personal congruence in all mass investiga-

tion, are some of the influences which tend artificially to lower the size and significance of correlations. Furthermore, though single coefficients may not seem high, the frequency with which moderate coefficients are obtained (as in our own results) is itself important. The discovery of *trends* and not perfect prediction must at the present time satisfy investigators who are studying the organization of personality.

Theories of Generalized Psychomotor Functions. (*a*) *Dodge's "Conjecture."* Reflecting upon the significance of his research on refractory phase, Dodge (*36*) raises directly the problem of general *vs.* specific functioning of the nervous system. The published statement of his hypothesis is very brief. He emphasizes the fact that "the final common path is known to be constantly subjected to impulses from various neural levels, so that any response originating at a reflex level is superposed on a complex dynamic background of pre-reflex conditions" (*36* p. 102). He also speaks of the greater retentive and dominating character of general attitudes as compared with specific (*36* pp. 21 ff.). Bringing these two ideas together Dodge concludes that specific ideas or acts are constantly influenced by the more persistent and more general mental attitudes or motor dispositions. The significance of this theory for psychomotor consistency is indicated in the following quotation:

> If we further postulate a relatively slowly changing group of emphasized R's which are regularly produced with every new R, must we not therewith grant our tissue a kind of personality? (*36* p. 144.)

According to Dodge's theory the genetic sequence of events would be somewhat as follows: Adjustments on the reflex or simple S—R level are unified into complex dispositions. This unification proceeds according to the principle of "persistent systematization" (*36* p. 131),

and this principle (akin to *Gestaltung*) is a basic property of the nervous system, "a precondition not a consequence of association." Now these newly formed motor dispositions are no longer clear-cut habits of action; they have lost their specific quality; and are to be regarded as blended, fused, perseverating tendencies to action, whose threshold of arousal is low. When specific stimulation occurs some specific adjustment is called forth, but also these general dispositions are inevitably aroused. The eventual action follows, accordingly, a common path created by the specific stimulus, the specific determinants, *and* the generalized determinants.

To illustrate with an example from our experiments: The subject is asked to look at a line 18 inches from the edge of a table, then to close his eyes, and to slide his hand out until the tip of his middle finger seems to him to touch the line. These directions and the visual cue arouse the specific movement required for the performance of the task; but they also arouse some motor disposition intricately developed from previous movements, from temperament, and from other internal conditions. The subject, we will say, moves his hand too far. This error is not an error solely in the estimation of 18 inches, for he overproduces most of the distances assigned him. The error in the "18 inch habit" shows a correspondence with the error in the 8, 12, 15, and 20 inch "habits." Furthermore, the same subject is very likely, according to our results, to show centrifugal exaggeration of movement in other experiments. We have here, then, strong evidence for Dodge's assumption that *persistent motor dispositions* are aroused in such a way that they can exert an influence on widely diversified habits of movement. To use his terminology we must "grant our tissue a kind of personality."

(*b*) *Other Suggestions.* Hsiao's theory of learning (*72*) is indirectly helpful. From his point of view both perception of the situation and the overt response to it start from a common "need." There is first an "irradiation" of the need, and the direction of this irradiation is determinative for both the perception and the response. Applied to our problem this theory would regard such generalized psychomotor functions as we have discovered as dynamic tendencies or needs, influential in determining both the attitude toward the task and the response. This point of view is closely allied to Lloyd Morgan's conviction (*112*) that the "higher" diffuse dispositions determine the "go" of specific response, and to Woodworth's theory that integrated habits may turn into drives (*201*). It is a thoroughgoing dynamic doctrine, and although not explicit in its account of the manner in which irradiation operates, is, so far as it goes, in agreement with our results.

These dynamic points of view bring to mind Köhler's application of his theory of "direct dynamical determination" to the field of movement. He writes,

> It seems to us that the nervous system sends to the muscles such innervations,—"pilots" them in such a way,—that the activity produced diminishes the tensions which exist each time in the cerebral field between the self and its environment. If that is correct, behavior, observed objectively, can serve in a large measure, as an indicator of the forces and tensions which in the interior, that is to say, in the cerebral field, tend to modify and, as a rule, to ameliorate the relation between the self and the environment (*90* p. 389).

The "self" in this formula is, of course, regarded as composed of quite permanent dispositions, hence it is evident that Köhler like Dodge views behavior as a resultant produced both by the exigencies of the specific situation and the rather stubborn tensions and tendencies

of the personality. The latter factors, of course, account for the self-consistency of expressive movement.

Personalistic Theories. Under this heading we may classify all those theories which hold that problems of expression cannot be treated separately from problems of the self. According to these theories movement is regarded as consistent because personality is consistent. As the *Introduction* has shown, this simple reasoning involves a surprising number of problems. Specifically, one would wish to know to what extent the inner personality itself is consistent, to what degree movement is expressive of personality, and whether movement is consistent in the same sense that inner traits are consistent. This last question can be made clearer by asking whether expansive and centrifugal motor tendencies, for example, have exact counterparts in the inner nature of a man, and whether such inner traits as insight (self-knowledge) and social intelligence have exact motor counterparts. All these questions are puzzling and have not as yet received satisfactory answers in personalistic theory.

Personalistic theories hold that most movements, even the simple adaptive responses which serve the temporary ends of volition, show more than momentary purpose: they show personality. Stern is the outstanding representative of this school, and the following quotation (somewhat freely translated) expresses his view.

> Interpretations of symptomatic acts bring us directly to the "Person," especially to the sub-wholes within the Person, to *Gestalten*, states, dispositions, phases, types, abnormalities, etc. The total Person (a terminal concept) simply determines the direction of the acts.
>
> But this directedness is the decisive thing for all interpretation of the significance of personal acts. *The direction leads unconditionally and unexceptionally from the single factor to the totality*, for example . . . from one characteristic of handwriting to an emotional quality, and from this to the structure of character.

> This direction toward totality is at the same time a progression from outside to inside, from surface to depth, from actuality to disposition.
>
> It is methodologically of the greatest importance to distinguish this characteristic course of experience from all others in which single factors are related to other single factors. Let us call the relationships in which significance is sought in the depths of personality, "vertical relationships" (the term surely needs no apology); then opposed to this method are others which may be said to deal with "horizontal relationships."
>
> For decades scientific work with the "Person" preferred to treat "horizontal" connections. Investigators analyzed and separated, determined functional, causal, and comparative relations, expressed the degree of co-variation by measures of correlation (*175* p. 73).

Stern then shows how superior is the method which seeks consistency in the "vertical relationship." He minimizes the study of correlations between single factors. Expressed in our terminology, he is critical of the value of investigating *correspondence,* and strongly prefers to stress the study of *congruence.*

Stern is alert to the genetic problems raised by his theory, for example, the question as to how the adult Person acquires functional unity in the face of the diverse and discordant influences which operate in the course of his development. Stern's intricate and extensive treatment of the vital dialectic of integration and differentiation (*172* Ch. 13) attempts to solve this problem. In his theory not all acts are equally expressive, for not all are equally bound to the personality (*173* pp. 162 f.). Originally many movements are simply adaptive and not expressive; with repetition they take on more and more of the inner qualities of the Person; and later, if the repetition continues, a point is finally reached where their expressive significance begins to decrease. Behavior which was formerly organic becomes more and

more mechanized and stereotyped and no longer reflects the vital patterns of the total personality. Conventional modes of greeting and manners of speech, for example, may be quite without expressive significance.

Having Stern's *Personalistik* in mind it would not be difficult to pass to H. Gross's theory that "every individual quality is merely a symptom of a whole nature, and can be expressed only by the whole complex" (*62*), or to Spranger's view that "the totality of mental life is expressed in every individual act" (*169*), or to Klages' "basic law" of expression (*86*) that each expressive bodily movement "actualizes" the tensions and drives of the personality. The details of these systems differ, but their personalistic intent, at least so far as expressive movement is concerned, is essentially the same. Klages is particularly rich in suggestions. His theory maintains that there is correspondence in the different provinces of movement: speech, handwriting, and facial expression. They have a common *Formniveau.* Yet at the same time their interpretation depends upon their personal reference. For example, the size, breadth, speed, and angularity of handwriting have a special significance depending upon the general "rhythm" of the individual's movement. Furthermore, Klages like Stern, allows for the "displacement" which convention and culture produce in the normal course of expression. The effort of men, in spite of this displacement, to express their personalities consistently in their movement furnishes the problem for his interesting theory of the *Gestaltungskraft.*[1]

These personalistic theories have the distinction of defending the position of common sense. They also have

[1] A convenient summary of Klages' view is given by Stern (*174*) who considers him to be a "decided representative of Personalism."

the merit of adequacy. Whereas the theories of identical elements and unique traits are too narrow to embrace all the consistency which our experiments disclose, the personalistic conception readily adapts itself to the discovery of a high degree of consistency. In fact, the more consistency found, the better for personalistic theories. This advantage, however, may turn into an embarrassment when we ask why movement is not even more consistent than we find it to be? If personality is, as Stern maintains, an *in sich geschlossene Ganzheit,* why is not movement more perfectly self-consistent? Why, also, should we secure quite decided evidence of, say, an "emphatic" factor in movement which does not immediately relate itself to some recognized trait of personality, and why is it that certain inner traits and interests have no visible counterpart in the general qualities of movement? In short, if movement derives its consistency from the consistency of the self, why should it be so difficult at times to trace this dependence?

Such questions signify that it is not easy to apply personalistic theories concretely to our data. Thus far Personalism has rested largely upon *a priori* reasoning and dialectic. Little attempt has been made to relate experimental evidence of consistency to these theories of the total personality. There is, however, no reason why, with the aid of experimentation, an inductive and empirical personalism should not some day be written.

Summary Statement. Our experimentation has been confined to peripheral motor consistency, but our search for a theory has necessarily led us to consider various views which embrace central as well as peripheral organization. A summary statement of our criticisms and evaluations of these views will perhaps serve as a

scaffolding for a more adequate, but as yet unconstructed, theory.

Taken literally the theory of identical elements does not help us. The correspondence of our measures does not seem to depend directly upon the recurrence of common features in the stimulus situations or in the use of identical nerve and muscle paths. It is, of course, not to be denied that any two measures that intercorrelate may do so because of the very complex similar *effects* which the tests have upon the subjects. We may say, for example, that voice intensity and point pressure correspond ($r = +.57$) because the two very different stimulus situations arouse a common tendency to make an emphatic response. But such a complex "common tendency" is equivalent to the admission of a general *disposition* to make emphatic responses. And a general disposition cannot literally be subsumed under the traditional concept of "identical elements."

Nor is the doctrine of "unique traits" easily reconciled with our results. In the first place the doctrine usually implies that traits are not only unique but also specific. Our results on the contrary show correspondence in broad clusters of paired, composite, and grouped measures. Far from being specific our motor traits are distinctly broad. Furthermore, absolute independence between even these broad clusters is not the rule. The uniqueness which is found in statistical studies is often an artifact: measures taken from an entire population are grouped by the investigator in such a way as to *construct* relatively independent variables. That the resulting statistical conception of trait is different from the clinical conception is clear. Unique statistical traits are not necessarily unique personal traits. Finally, we have to criticize the particular application of the theory that holds

bodily traits to be independent of mental. This conclusion is not only opposed to common sense, but is based on experiments which are both narrow and naïve.

Several theories assume that there are processes of integration, generalization, and spread, which result in complex psychosomatic dispositions that exert a dynamic determination upon specific movements. Variations of this view advanced by Dodge, Hsiao, and Köhler have been briefly described. Although the application of these dynamic theories to expressive movement has not been attempted in detail, their direction seems to be in line with our results.

Finally, the fundamental question of the relation of movement to the total personality is raised by personalistic theories. They assume that there is a unity in the Person, and that this unity necessarily entails consistency between expressive acts. The theories account almost too well for our results, since they seem to demand congruence in nearly every response, as well as a striking degree of correspondence between single measures. Certainly both congruence and correspondence are revealed in our data, but we do not discover the almost perfect consistencies that the personalistic theories at present seem to require. This criticism, however, may be no more than a reflection upon the adequacy of our own methods; or it may betray the absence of an experimental basis in the personalistic theories as they stand. If the latter interpretation is correct, there is no reason to suppose that these theories should not eventually be capable of adapting themselves to experimental findings. *Fundamentally our results lend support to the personalistic contentions that there is some degree of unity in personality, that this unity is reflected in expression, and that, for this reason, acts and habits of expression show a certain consistency among themselves.*

These various considerations are advanced as contributions toward an adequate theory of the consistency of expressive movement. It would be unwise at present to attempt to make a more final statement. We hold only that our evidence, so far as it goes, favors an hypothesis that is essentially dynamic and personalistic rather than one that is essentially static or specificistic.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

The Problem. It is a matter of common belief that the gestures or expressive movements of a person are consistent with one another. A man's style of writing is supposed to be harmonious with his manner of speaking, his facial movements with his posture, and his gesticulation with his gait. This belief undoubtedly derives from an underlying conviction that all the mobile features of the body are avenues for the expression of personality.[1] If personality is self-consistent, it is reasoned, its expressions must in turn be consistent among themselves.

Two quite distinct issues are involved in this common sense view. The *first* concerns the assumption that an individual is essentially self-consistent in his personality, and that this self-consistency of personality is reflected in some direct and uniform fashion in his actions. This assumption is the foundation of all practical attempts to diagnose personality from external movement. But since so little is known concerning the reliability of expressive acts themselves, this first issue is scientifically premature. A *second* problem, namely the self-consistency of expressive acts, is logically prior to the first, and is at the same time a far more accessible problem. All the investigator needs is a sufficiently large sample of records of normal movements from a sufficiently large group of subjects, together with statistical tools for the study of the variability of his measures. In all the previous

[1] "With the exception of hard or dense tissues (bones, cartilage, tendons)," writes Boven, "the activity of character is seen manifested in the functioning of all the tissues of the organism" (*19* p. 119).

literature on individual differences there are very few references to this simple and obvious problem of intra-individual consistency in normal movement.

While pursuing this neglected subject with instrumentation and statistical tools, the investigator should not lose his perspective. He must remember that he is working on only one of the three possible avenues of approach to the problem of the organization of personality. He is not touching directly the riddle of integration in terms of the internal structuration of the dispositions, traits, attitudes, or interests of the personality; nor is he attacking the equally important *social* aspect of expression, namely, its effect on other people in terms of judgment, interpretation, and reputation.

The investigator must be careful also lest in his single-minded pursuit of the objective correspondences among his measures, he overlook non-quantitative evidence of motor consistency. If he works with a broad view of his problem, he will, in fact, quickly discover that in addition to the mathematical correspondence of his measures, he has before him less tangible but no less arresting evidence that consistency of expressive activity lies not only in the correlation of his measures, but in their congruence (meaningful interrelation) as well.

The Method. A full set of experimental records was obtained from a diversified group of 25 male subjects. In three private experimental sessions with each subject, over thirty simple tests were administered. Most of these tests yielded several records (*e.g.*, speed, size, and pressure of movement); many of them were performed with the right and left arms or legs; and most of them were repeated in order to determine their reliability. As a result, some three hundred original measures were obtained in all. In the treatment of the data the measures

were pooled in various combinations in order to obtain the most reliable and most meaningful groupings. The experiments were supplemented by ratings, gathered both from within the laboratory and outside, concerning aspects of movement which were not easily accessible to experimentation. The total amount of data at our disposal is, therefore, quite extensive.

The tests themselves were devised to provide accurate records of the subjects' natural and normal manner of performing simple tasks. Usually no special instruction regarding rate, pressure, size, or style of behavior needed to be given; when necessary the instruction included such phrases as, "at a convenient speed," or "in your usual way." The simplicity of the tasks (walking, reaching, counting, tapping, writing, drawing simple figures, and the like) helped to secure naturalness of performance. The only instrumentation of an elaborate order was a pressure board, stylus, and apparatus for studying muscular tension. But even with these instruments the subject had to perform only the very simplest of tasks and was free to choose his speed and style of approach.

Although the experiments themselves were purposely constructed to be as natural as possible, the problem of treating the voluminous records is unavoidably complex. Certain standard statistical aids are at our disposal. Among those most frequently used are Spearman's formula for the correlation of rank orders, and his method of sums for computing the correlations between two composites of tests. In dealing with a large number of measures of the same variable, or combination of variables, the first two of Kelley's formulae for average intercorrelation are used. The technique of building combined and composite measures from a large number of interrelated original measures entails the application of the

Spearman-Brown prophecy formula. All these aids, however, must be applied with strict caution and with due consideration for their limitations. As a matter of convention, the usual corrections for reliability and internal consistency are applied, even though they do not in themselves advance the psychological significance of our results. Raw, uncorrected coefficients are always given in the text, and it is upon these that most of the conclusions are based. We have attempted to map out a broad and almost unexplored territory, and believe that the indications obtained with simpler statistical methods are all that are at present justified.

Experimental Results. Our argument is that wherever satisfactory reliability or internal consistency is found among our measures we have *prima facie* evidence of consistency or integration among the expressive acts of our subjects. In every case, however, this evidence is examined to determine whether it is psychologically intelligible as well as statistically sound.

(1) The average of the repeat-reliabilities of the separate tests is +.684. In view of the fact that this figure is uncorrected, and that the median time devoted to the tests from which it is derived is only 30 seconds, we are justified in stating that a subject repeats on later occasions a certain act in essentially the same way. Specific performances are to a rather high degree constant.

(2) Tests repeated in the same experimental session have average reliabilities of +.75; those at different sessions, +.644. It appears therefore that temporary factors (*e.g.*, mood) play an appreciable but not exclusive part in determining the constancy of performance.

(3) The average correlations for tasks performed with different groups of muscles are just about as high as those performed with identical muscles. This finding confirms

many studies on cross-transference, and proves that the performance of single tasks is not specific to single members of the body. We have here clear evidence for inter-muscle consistency.

(4) Studying all the 14 measures of speed, we find no general speed factor or uniform "psychic tempo" among our subjects. There is, on the other hand, evidence that each speed measure is itself reliable, indicating a high degree of constancy in "specific speeds"; in addition, there are three somewhat broad speed factors, namely *verbal*, *drawing*, and *rhythmic*.

(5) Seventeen composite measures were constructed out of several single measures (an average of 5.89 single measures in each composite) which intercorrelated too highly to be considered independent (average r, corrected, $= +.838$). The possibility of constructing these composite measures is itself evidence of consistency of a complex order. To illustrate, the relative area of a subject's writing is very constant, whether measured by his copying a paragraph, writing on a pressure board, with a crayon on paper, on the blackboard, or in a box of sand with his foot. We have no choice but to make one composite measure out of all these tests. They indicate conclusively that "area of writing" is a highly consistent expression among our subjects.

(6) We are left, after an analysis of all our data, with thirty-four variables (that is, reasonably distinct paired, combined, or composite measures). The *Appendix* gives the intercorrelation of these variables and their reliabilities. The reliabilities (corrected) average +.813. Here again we find that the constituent items in a constructed variable have a decided tendency to correspond, a fact which further indicates the marked spread among our measures.

(7) On the other hand, we find little evidence for a "general motility factor" saturating all the measures. The median coefficient for all variables correlated with one another is only +.05. Corrections for this figure enhance it enormously, but probably to a large degree, spuriously. On the whole it is undoubtedly unsound to claim that our results disclose a general factor of "energy" or psychomotor "power."

(8) Examining still more closely the table in the *Appendix*, we find that the lack of evidence for a general factor is offset by the promise of the discovery of group factors. By clustering the variables in such a way as to obtain the highest possible average intercorrelation and internal consistency, factors are disclosed that are not only self-consistent and relatively independent, but are also psychologically meaningful. This procedure results in the isolation of an *areal* (or expansive) factor, a *centrifugal* factor, and a factor of *emphasis* in movement.

(9) The discovery of such motor factors or "traits" in our population of subjects suggests that psychological and not physical categories are fundamental in the study of movement. *Emphasis*, for example, partakes, to be sure, of the physical measurements of pressure, but it involves other tendencies, such as the disposition to fill only a small part of the space at one's disposal in writing. The spatial dimension of movement (homogeneous enough from the point of view of physical measurement) breaks down into two almost independent psychological dispositions, the *areal* and the *centrifugal*. Speed apparently is another factor that is homogeneous only to physics; in our results speed seems to split into three relatively independent rates of movement. Many of the speed measures correlate more highly with non-speed measures than with each other. In short, physical categories of

movement are unsuitable models for the *psychological* study of expression.

(10) There remains the question of the variability of a subject's scores treated by the method of Pearson's coefficient of variation. Examining the variability of all the subjects on 11 representative composite measures, we discover no evidence for a general factor of individual consistency. The average inter-*r* of all the variability scores is only +.02. By this method it does not appear that a given subject tends to be equally consistent or variable throughout all the tasks.

(11) There is, however, a strong suggestion of the presence of two independent composite consistencies. The variabilities of the three most complex grouped measures (the three "group factors") correlate on the average +.403 (corrected, +.67). Similarly, the composite measures of a more specific order correlate positively, with an average coefficient of +.161 (corrected, +.53). These two composite variabilities (the general and the specific) are quite independent, their average inter-*r* being −.14. These data, though inconclusive, suggest that variability may be a somewhat consistent trait at a lower or at a higher level of integration, but not at both levels.

The extreme indirectness of this method of obtaining light on intra-individual consistency tempts one to abandon the guidance of statistics at this point, and to look directly at the data which are available for the individual subject, in order to obtain, as it were, a complete psychomotor portrait. This method of studying simultaneously *all* records for an individual subject brings us directly to the discovery of psychomotor *congruence.*

(12) Four case studies, selected for their diversity, are presented. Something of the professional and social life of the subjects and of the impression they make on others

is given. With this approach we are forced to the striking conclusion that virtually no measurements contradict the subjective impression of the personalities. The measures faithfully record what common sense indicates. Even measures which do not correspond statistically fit into the picture in such a way as to be readily intelligible and psychologically congruent.

(13) It is shown furthermore that Dr. Saudek's analyses of the motor qualities of these subjects, based exclusively upon their handwriting, agree closely with the results of the experiments. Chirography yields essentially what our instruments of experiment yield.

(14) A group of judges is able to match with considerable success the kymographic curves and the handwriting obtained from these four subjects with brief thumb-nail sketches of their personalities.

Theories of Expressive Movement. The problem of the organization of an individual's habits of movement is clearly not a simple one. On the one hand there seems to be no general psychomotor factor which directly and unambiguously determines each and every act. If there were we should find it possible to assign to each subject a single score on motility, and expressive movement would become simply a homogeneous quantitative variable in personality. On the other hand, we find it impossible to conclude that acts are completely specific, determined only by independent habits or by the individual situation in which they occur. There is obviously neither complete generality nor complete specificity.

Our results favor an hypothesis that there are organized psychomotor dispositions or expressive traits. The fact that certain of these traits are discovered in an analysis of data obtained from an entire group indicates that they are to a degree universal and scalable. The fact,

however, that these same traits (and others besides) are especially prominent when one undertakes a clinical study of individual cases, indicates that organization has likewise its strictly individual aspects. Some motor traits are scalable, others are peculiar to the individual.

It is shown that the theory of identical elements, which is essentially an interpretation of psychomotor organization in terms of the specificity of functioning, is untenable. The amount of "transfer" among our measures is not proportional to the number of identical elements contained in the various series. Especially do our results call into question the notion that there is no functional unity between mind and body.

More promising are the theories of the dynamic interrelation between the cerebral field and the peripheral field of movement. In such theories the higher centers reflect their own organization in consistent manners of expression. We find instances of this line of reasoning in the writings of several contemporary psychologists. Finally, we are faced with the position of personalistic psychologists who regard movement as lock-stitched into the fabric of personality. Their work, however, has been almost entirely deductive; ours is inductive. We do not find it easy to say that the aspects of expressive movement which we have studied are direct and unambiguous reflections of "inner" disposition. This problem is one in psychodiagnostics and lies beyond the scope of our inquiry. Nor do we find the almost perfect consistency in movement which certain statements of the personalistic theory would seem to require.

With the broader aspects of this theory, however, our results are in agreement. Motor acts are not so specific as to be meaningless, and being organized they must reflect to a large degree the organization of the total brain

field. There are degrees of unity in movement, just as there are degrees of unity in mental life and in personality. It is surely not unreasonable to assume that insofar as personality is organized, expressive movement is harmonious and self-consistent, and insofar as personality is unintegrated, expressive movement is self-contradictory.

PART B

HANDWRITING AND PERSONALITY

CHAPTER IX

THE PRESENT STATUS OF EXPERIMENTAL GRAPHOLOGY [1]

Is There a Case for Graphology? The prevailing attitude toward graphology among American psychologists is one of skepticism and distrust. It is customary to find graphology contemptuously dismissed along with phrenology and palmistry, or deplored as a public menace of the "gold brick" variety. In the minds of most writers the commercial abuses in graphology seem to disqualify it for impartial experimental study. As a consequence of this point of view very few psychologists in this country have studied its methods or claims seriously; in their criticisms they seldom refer to more than one or two inconclusive experiments. The attitude of J. B. Watson (*193*) is typical. Leaning, as nearly all American critics of graphology do, upon the single experiment of Hull and Montgomery (*75*), Watson concludes that all claims for graphology are a "tissue of exaggerations" which "will not bear critical experimental testing." Unfortunately the investigation by Hull and Montgomery, considered so crucial, proceeds from a conception of graphology which is remarkably naïve and largely irrelevant.[2] Symonds (*178*) after conventionally citing this same study concludes that the average correlation obtained by Hull and Montgomery, −.016, "represents about the amount

[1] The writers wish to acknowledge the assistance of Professor Edwin Powers in preparing this review. It should be regarded as expressing the dominant trends in the field up to the middle of 1931. Since the review was written, interest in the subject, especially in Europe, has grown rapidly, and the output of new literature is striking both for its quantity and its quality.

[2] See pp. 192 f.

of assurance that one should give to the claims of graphologists." Since this famous experiment has only remote relation to most of the claims of graphologists and practically none at all to their methods of work, and since a vast amount of contradictory evidence is now available, it is a meager and insecure foundation for such a sweeping opinion.

In contrast to skepticism and neglect in this country we find an alert and sympathetic interest in Europe. In Germany, France, and Holland there are several societies and journals devoted wholly or partly to psychological graphology.[1] Many reputable psychologists on the continent are following the topic and experimenting in the field of handwriting. Its popularity is evidenced by the recent republication of the standard works of Preyer (*135*), Meyer (*109*), and Klages (*86*). Arnheim (*7*) and Theiss (*179*) have shown graphology to be a happy hunting ground for Gestalt theory, and confidence in the diagnostic uses of handwriting is illustrated by extensive experiments in the field of psychotechnics.

In brief it might be said that American scientists regard handwriting as unrelated to the deep-lying central factors of personality, and as a product essentially of peripheral manual movement; they also consider it to be influenced greatly by external conditions of instruction and example. Continental psychologists on the other hand see in graphic movement the quintessence of expression. It is a "crystallized" form of gesture, an intricate but accessible prism which reflects many, if not all, of the inner consistencies of personality.

In order to justify their confidence in graphology its champions advance many arguments. For example, it is well known that adults who have been taught identical

[1] The first serious American periodical in this field is now appearing under the title *Character and Personality*, R. Saudek, editor, Duke University Press.

systems of penmanship invariably develop some degree of individuality in their writing. That this graphic individuality is a stable characteristic is shown by experts, such as Osborn (*126*), who demonstrate that it is practically impossible either to disguise one's own handwriting completely, or to copy the handwriting of another person beyond the range of detection. Additional evidence for the deep-lying determination of graphic habits is found in the numerous instances where a writer has lost his right hand and has been forced to learn anew with his left hand, or even with his feet or mouth. Saudek shows irrefutably (*151*) that whatever the muscle group employed, an individual's scripts are fundamentally identical; they are all one and the same "brain-writing." Preyer's experiments (*135*), in which right-handed adults were trained to write with their left hand, lead to the same conclusions. Land (*95*) holds that such facts as these are convincing evidence for the superiority of script over tests as indices of personality. Handwriting provides material that is less artificial than tests and more convenient for analysis; and since it can be studied at leisure, it is superior to facial expression, gesture, and gait which are so fleeting and difficult to record.

Still more spectacular evidence comes from experiments with hypnotized patients. It is said that a patient who under hypnosis assumes the rôle of, say, Napoleon, tends in his writing to take on the characteristics of Napoleon's script. Saudek claims that this phenomenon has been established by many experiments, though Downey (*39*) is rather dubious of their worth. Mühl (*115*) describes the varieties of handwriting in one case of multiple personality, and Melcher (*105*) reports that he has found differences between the handwritings of different co-personalities, and also among individuals in their normal

state and under the influence of alcohol. Downey (*39*) summarizes certain studies on changes of handwriting in manic-depressives and other disordered personalities. In general it does seem that variations in personality are accompanied by variations in script, although our knowledge of this phenomenon is incomplete.

No one who has considered carefully the experimental and theoretical work on handwriting seems to deny the *a priori* case for graphology. Graphic movement apparently is not activity that is dissociated from the complexities of personality; on the contrary it seems to be intricately woven with the deep-lying determinants of conduct. Graphic movement is, therefore, expressive movement; hence its place in the present volume. If our thesis is correct that expressive acts are not specific and unrelated to one another, then handwriting itself will show consistencies with other expressions of personality. That this is indeed the case will be seen from the following review of literature, and from the experiments reported in succeeding chapters.

Methods of Graphology. There is a common misapprehension among psychologists and laymen to the effect that graphologists all work with particular "signs" or details of script which they hold to correspond invariably to specific traits, abilities, or interests. It is true that Michon, the founder of modern graphology, evolved an elaborate system of such signs, and that many of the more superficial graphologists continue to beguile the public with this method. As Downey points out (*39*) many of the correspondences which they claim appear to have little more basis than the sympathetic analogies of homoeopathic magic. For instance, long upper projections are identified with idealistic thought, long downward projections with earthly, material interests. But

the tendency among more thoughtful writers on graphology is to depart from these elementary theories, very much as theories of associationism are gradually yielding to the point of view of *Gestalt*. Crépieux-Jamin (*32*) writes that the study of details of script is to graphology as the study of the alphabet is to the reading of prose; hence his theory of "resultants," or interplay of many signs. Handwriting, he believes, is a "total harmony," to which each note (sign) contributes in varying degrees and with varying significance. Other graphologists go further, basing their conclusions on the recurrence of more general features of script, and upon their balance, the total impression or *Formniveau* as it is called. This relational type of graphology is well illustrated by the writings of Biäsch (*10, 11*). Most of the experiments devised by skeptical psychologists, who seek correlations between specific details of script and specific traits, are foredoomed to negative results. If we are to attain the most adequate validation, *the script as a whole and the personality as a whole must somehow be compared.*

Graphological methods at their best may be studied in the writings of Saudek. Having set himself "the task of accepting *none* of the graphological doctrines hitherto accepted which cannot be *proved* either by systematic experiment, or by statistical inquiry on the broadest basis," Saudek has made discriminating use of the microscope, pressure board, and moving picture camera. His methods have been described not only in his numerous publications but likewise by Brooks (*22*) and Meloun (*106*). Curiously enough Saudek is attacked in Germany as being too statistically-minded, and in America for his "vague speculations." The Germans, influenced so profoundly by Klages, regard handwriting as capable of revealing to the initiated the "essence" of personality;

but it is only by intuition and not by analytical or statistical techniques that the essence is grasped. American critics, on the other hand, with their leaning toward specificity and objectivity, regard all methods excepting the quantitative with suspicion. Saudek who takes a middle position satisfies neither school completely; but unlike most graphologists, he answers many of the objections raised by critics.

He admits, for example, that peripheral and external factors are of great importance in handwriting. Before analysis can start it is necessary to take into account the nature of the ink and of the writing surfaces, the pen and paper which have been used, and also their positions relative to the hand of the writer. The style of penmanship which the writer was originally taught must be considered next. Saudek's diagnoses differ markedly according to the nationality of the writer (*153*) also according to his age (*155*). Children and adolescents seldom show graphic maturity simply because their script remains closer in form to their copy book; but as they mature no two individuals retain identical scripts.

Besides graphic immaturity there are obstacles to spontaneous writing such as the conventionalization of clerical, official, or calligraphic hands. Conscious attention artificializes the script; and hence, according to Saudek, "the value of a handwriting for research material is the greater the more the attention is distracted from the writing." Conditions should always be the best possible for permitting free play to the spontaneous determinants.

Above all, the speed of writing must be taken into account, for speed, significant in itself, also modifies every other feature of writing and conditions every graphological deduction. Saudek (*150*) lists a number of criteria for determining the speed of writing from the script;

these comprise 8 primary, 4 secondary, and 5 equivocal symptoms. The primary symptoms include such criteria as increasing the size of the left-hand margin toward the bottom of the page, dotting the *i*'s ahead of the *i*-stems, writing the dots as accents, rising alignment, linking of words. All scripts show some rapid and some slow symptoms, but Saudek maintains that a preponderance of two primary, together with some of the secondary and equivocal symptoms, are significant of truly rapid writing.[1]

There is, of course, considerable variation in methods employed by graphologists. Some experts, notably Saudek, incline to base their interpretations on a scientific analysis of the numerous factors involved in handwriting. Others rely chiefly on general impressions and *Formniveau*. A few of the more popular practitioners still work exclusively with "signs," and like commercial physiognomists seem to have a small margin of success in spite of their questionable systems. Whether graphologists tend to agree in their interpretations regardless of the dissimilarity of systems employed, is a question that has interested Schorn (*159*) and Powers (*133*). Each of these investigators studied the agreement among five professional graphologists by comparing point by point their analyses of a single handwriting. On most of the points there seems to be rather high accord. Powers found only 3 definite contradictions between any two graphologists, as against 68 fairly certain agreements, and 4 points which were mentioned only by single graphologists.

One of the common criticisms made against graphologists is that they judge not so much by the script as by the

[1] Saudek does not seem to have proved that his judgments, derived from these signs, actually correlate with the measured rapidity of writing. When Downey estimated the speed of writing by inspection (probably employing less adequate criteria than Saudek's) she found a correlation of +.55 with measured speed (*39*).

material which is written. In their ordinary commercial practice they are, of course, entitled to make every possible use of the content of the writing, for their practical success in "reading character" and not their method is at issue. In strict scientific research, however, it is absolutely necessary to control this factor by the elimination of cues derived from content or style of expression.

Validation of Graphological Diagnoses by Psychometric Methods. As has been pointed out, the work of Hull and Montgomery (*75*) is cited enthusiastically by critics of graphology. These investigators obtained copies of a passage from 17 student subjects, each subject using his own pen and uniform stationery. These specimens were measured microscopically for some of the elemental "signs" which play a part in the systems of certain graphologists. The signs included fineness of line, lateral narrowness of *m*'s and *n*'s, upward slanting alignment, length and heaviness of *t*-bars. Each of the subjects, fraternity members, was ranked by the others on various traits such as ambition, bashfulness, forcefulness, perseverance, and reserve, with which these "signs" were supposed to correspond. The average correlation between ten such sets of measurements and ratings was $-.016$, the range from $-.45$ to $+.38$. The most obvious sources of error in this "test" of graphology are the following:

(1) Practically no graphologist at the present time, certainly not the leading authorities, claim any fixed and invariable significance for any of these "signs." Although borrowing in part from the signs listed by Crépieux-Jamin, the investigators omitted entirely his doctrine of "resultants." No attempt was made to balance and synthesize the signs or to work with the form-quality of the whole script. Probably the only one of these signs to which Saudek would give much credence is the width of the

m's and *n*'s, which incidentally showed the highest positive correlation with any of the traits studied.

(2) Nor would any graphologist assert that there is a proportional correspondence between the gross magnitude of a sign and the magnitude of a trait.

(3) The criteria by which the graphological diagnoses were tested were ratings by fraternity members on trait-names. Ratings of this sort are known to be unreliable. Furthermore, the significance of the trait-names to the raters was not necessarily the significance intended by the graphologists who originally employed them.

(4) The subjects were an immature and homogeneous group making for an unreasonable restriction in variability. Also their script, being a copy, lacked spontaneity.

The deficiencies in this method are so patent that one is neither surprised nor convinced by the negative results. The study bears virtually no relation to the methods employed in graphological practice, nor to any of the current theories of the relation between handwriting *as a whole* and personality *as a whole.*

A somewhat similar experiment by Brown is reported by Hull (*74*). She, however, used spontaneous writing rather than copied material. Thirty sorority women ranked one another, and the same fragmentary, microscopic measures of their scripts were made. Neatness and individuality of handwriting were also rated by 10 judges. Brown found a slight tendency (+.23) for upward sloping alignment to correlate with ratings on ambition; graphic and personal neatness correlated +.23; and graphic and personal individuality correlated +.15. Most of the coefficients were around zero. Another widely quoted experiment with essentially negative results is that of Omwake (*125*) who examined the ability of untrained judges to determine intelligence from handwriting.

These investigations suffer from many of the same deficiencies as that of Hull and Montgomery.

A different type of approach which yields strikingly different results is Land's (*95*). This study eliminated ratings as a criterion, and employed only the general graphic features of slant and alignment. From among 200 students (aged 17 to 50 years) Land selected 10 whose normal scripts showed the most extreme backhand, 10 whose general alignment was most downhill, and 10 controls who possessed the usual forward slant and uphill alignment. These subjects were given the Pressey X-O test. The scores on Part I ("Worries"), Part III ("Things thought wrong"), and Part IV ("Unpleasant things"), for the three groups were as follows.

PRESSEY SCORES IN CONTRASTING GROUPS OF WRITERS (after Land)

TYPE OF EMOTIONAL RESPONSE	BACKHAND GROUP	DOWNHILL GROUP	CONTROL GROUP
Worries: average	45.7	47.8	19.0
range	(32–60)	(30–73)	(3–27)
Wrongs: average	72.9	73.3	63.8
range	(46–100)	(61–100)	(42–81)
Unpleasant: average	44.2	44.5	30.6
range	(26–59)	(29–63)	(8–52)
Total average	162.8	165.6	113.4

By this test, then, the two selected groups are proved more "emotional" than the control group. In spite of overlapping the differences are significant. From Land's tables may be calculated the Difference / S.D. between the sum of the three tests for backhand and control groups; the result is 4.3; for the downhill and control groups, 4.2. Both differences, therefore, are much superior to chance.

Downey (*39*) also attempted to test several claims of

graphologists, but did not confine herself solely to the measurement of specific signs. She asked 12 psychologists to rate 28 to 30 other psychologists, whom they knew moderately well, on several traits. By analyzing the handwritings of the subjects, she classified them for the same traits.

SOME CORRELATIONS BETWEEN GRAPHOLOGICAL RATINGS AND RATINGS BY COLLEAGUES

TRAIT RATED	GRAPHOLOGICAL SYMPTOMS ON WHICH DOWNEY BASED HER JUDGMENTS	CORRELATION
Attention to Details	Smallness of letters, regularity and invariability, even spacing	+.61 ±.08
Pride, or Feeling of Self-Worth	Size of letters and ornamentation of capitals	+.24
Explosive-Inhibited	Klages' criteria, such as spontaneity, exaggerated size, rapidity and pressure (hyperkinesis), and variability of script	+.53
Aggressiveness	Forceful strokes and pressure	+.23 ±.13
"Temperament"	Slant and fluctuations	+.27

Since both sets of ratings were on a five-point scale, 20% of the correct placements might have been obtained by chance; the first trait represents 48.2% correct placements, the last two represent 33.3% correct. Downey believes that the homogeneity of the group of subjects, and her partial reliance on graphic details rather than on the script as a whole, probably reduced the extent of agreement with ratings.

In Germany, graphology has been seriously applied to industrial and personnel investigation. Couvé (*31*), for instance, studied employees of the *Deutsche Reichsbahn.* The foreman rated 12 employees on a three-point scale,

according to their efficiency. The graphologist, judging their handwriting, agreed in his placements 10 times and was only one step removed in the other two. Of 108 further graphological ratings by nine untrained judges, 68% were correct, 16% were one step removed, and 16% were wrong (two places removed), where 33.3% would be right or wrong by chance. At the *Institut für Industrielle Psychotechnie*, von Kügelgen (*94*) analyzed the writing of 48 pupils, classifying them on a five-interval scale for nine traits. Ratings by teachers gave complete agreement for 66%, where about 20% would be correct by chance. Other investigators have employed objective, but not readily quantified, criteria. Mayer-Benz (*104*), for example, found considerable agreement between her interpretations of handwriting and the case histories of the writers as recorded in a mental clinic. In the industrial field, Seesemann (*164*) analyzed the handwritings of 20 mining officials, and obtained Yes and No answers from foremen or directors to 14 questions dealing with the traits of each writer. Using only this rough basis for comparison, Seesemann claims that 93% of the judgments were in agreement with his findings.

Henning (*67*) criticizes some of these experiments on the grounds of artificial selection of the subjects, and shows that among 2000 unselected children, there was no agreement between aptitude tests and ratings on proficiency as determined from handwriting. Von Foerster (*50*) had the scripts of 70 pupils rated by five judges according to their aptitude (*Gewandheit*). Though the judges agreed among themselves, correlations between the ratings and tests of intelligence and special ability reached only +.29 or less. He also measured 17 of the traditional graphological "signs," but (as was to be expected) he found that they failed to give significant correlations.

Psychometric methods of validation, it seems, give contradictory results. Detailed studies of the diagnostic significance of specific signs are uniformly negative. Most other methods, however, yield positive and sometimes striking results. This state of affairs means that conclusions from this type of research can be drawn only for the specific conditions under which the study is made. On the other hand, the results obtained by the following method of validation are of more immediate importance for the theory of expressive movement.

Validation by Sorting and Matching. Psychometric validation leans chiefly upon the measurement of specific "signs," ratings, correlations, and other precise, quantitative aids. A less artificial, but equally controlled, method is the comparison of graphological judgments, based on the script as a whole, with estimates of personality based on facts (as distinct from ratings) or on complete case histories.

One form of this method requires that two groups of scripts be shuffled together. Judges are then told to *sort* the scripts into the desired groups: male and female, old and young, or honest and dishonest. This was the method employed by Binet (*15*). It will be remembered that, while 50% success at sorting would be achieved by chance, actually from 61% to 92% success was obtained by Binet's various judges, the experienced graphologists being the most accurate. Thus, while imposing scientific controls, Binet allowed the judges to make free use of total impressions, or any form of analysis which they desired. It is further noteworthy that, as a reputable psychologist, he gave it as his opinion that there was some truth in graphology, that it was a "science of the future." He was fully sensible of the graphologists' limitations, since none of them achieved perfect scores,

but he pointed out that the results of all were superior to chance.

Downey (*38*) repeated Binet's experiment on the determination of sex by handwriting in 1910, also using envelope addresses as her material; but she controlled the factor of the sex of the person to whom they were addressed. Her 13 non-expert judges sorted 67.3% (range 60% to 77.5%) of her 200 specimens correctly. She amplified Binet's observation concerning the letter formations, and other graphic characteristics which are commonly associated with the two sexes.

Additional experiments on judging sex from script have been carried out by Kinder (*84*), Newhall (*122*), and by Broom, Thompson, and Bouton (*23*). It is doubtful whether these have added anything to Binet's and Downey's results, except slight improvements of technique. That the imperfect validity of these investigations should appear to some writers as a disproof of all graphological claims is curious, since, as Rice (*140*) points out, "from Crépieux-Jamin down to the present author, every graphologist has realized that sex simply does not exist in handwriting." Most graphologists indeed insist on being told the sex and age of the writers before they undertake an analysis. Crépieux-Jamin stated that "there will always be examples of effeminate men and virile women, the former will employ the feminine, and the latter a masculine type of handwriting. It is the psychological and not the physiological personality which is revealed in handwriting." Saudek's view is the same. From an analysis of many scripts of both sexes, Wittlich (*197*) concludes that the average man shows 64% of the typical masculine and 36% of the typical feminine "signs" in his writing and *vice versa*.

The most successful investigation to employ this tech-

nique of sorting is that of Saudek on honesty (*150*). It represents, moreover, his chief attempt to conform to the validatory standards of scientific psychology. He lists some 10 general "signs," any four of which, occurring together in a script, are said to reveal dishonesty in the writer. Chief of these is the curious feature of leaving *o*'s, *a*'s, *d*'s, etc., open at the *bottom*, a symptom which "seventy years of experience of all graphologists" have shown to occur only "in the handwriting of dishonest, or at least very insincere persons." He states that in the files of the French police, 30% of the signatures of "habitual thieves of both sexes" show this symptom. In his own experiment, he obtained specimens of 73 handwritings from 18 business firms, without knowing the writers, and diagnosed "dishonesty" in 14 cases. In all of these he was correct. In no case was an "honest" person diagnosed as dishonest, but one writer who showed only 3 of the 10 "signs" was classified as honest by Saudek and was called dishonest by the firm. Similar experiments by Saudek, giving equally good results, are summarized by Meloun (*107*). As striking as these experiments are, it is necessary to warn the reader that for various reasons some psychologists are critical of the concepts, procedure, and criteria employed (cf. Symonds, *178*, p. 526).

Like the method of sorting, the matching method employs large numbers of judges, but the group of subjects can seldom exceed 10 in number. The following are the chief varieties of matching.

(1) The graphologist makes anonymous personality sketches of the subjects (writers); judges who know the subjects personally then try to decide which of the sketches refers to each of the subjects. This method is employed in Chapter XI.

(2) Reversing the process, a psychologist or some other

judge makes personality sketches of the subjects, which are then given to the graphologists, who attempt to match the studies with the scripts. This method is employed in Chapter X.

(3) Trained or untrained judges may be shown the handwritings (which they have not previously seen) of people with whose personalities they are in a measure familiar. The matching or identification then takes place as usual. Arnheim's experiments are of this type (cf. pp. 13 f.). His judges matched their impressions of the personalities of Michelangelo, da Vinci, and Raphael with their scripts. He also used photographs of unknown persons, which had to be matched with their handwritings. Von Foerster (*50*) improved on this method by having eight writers appear before a group of 25 judges; the judges were required to guess which subject produced each specimen of script. One judge achieved 6 correct matchings; three were totally unsuccessful; and the average number of correct matchings for each judge was very slightly above chance.

It is somewhat unfair to confine the method to the matching of script with such partial and fragmentary information concerning the subjects. If graphology is to be allowed to justify its claim of revealing the total personality, the judges should have *all* the available data concerning the personalities of the writers. Wolff, as was shown in Chapter I, required friends of the subjects to identify case studies based on graphological analyses, allowing them in this way to employ their total knowledge of the writers. The most important experiment of this type is that of Bobertag (*16*), and his results, if they are confirmed by further experimentation, are the most striking so far obtained.

Bobertag used 6 professional graphologists, 5 writers

or subjects, and 15 judges. The writers and judges were teachers in the university and knew one another fairly well. The graphologists were first asked to make thorough sketches of the writers from their scripts. These 30 sketches were then given to the judges to identify. They were informed that each of the 5 writers was represented 6 times. Of the 450 matchings, 80.7% were correct, 4.7% were left undecided, and only 14.6% were incorrect. Theoretically one-fifth, or 20%, of the matchings in this experiment might be correct by chance. Powers (*133*) verified this probability by a statistical experiment, drawing numbered cards, under the conditions of Bobertag's study. He found that 19.77% were correct by chance, the highest proportion achieved by any one hypothetical "judge" being 9 out of 30, or 30%. Clearly, even the worst of Bobertag's judges obtained a result enormously superior to chance. Some personalities, Bobertag noted, were much more easily identified than others, and some graphologists wrote better sketches than others. But on the whole the sketches of the same subject by different graphologists seemed to be in substantial agreement (cf. Schorn and Powers, p. 191). Schrijver (*160*) summarizes and criticizes this experiment and one of the same general type by Powers, to be reported more fully in the following chapter. In a later article (*17*), Bobertag defends his application of scientific techniques as contrasted with the prevailing subjective methods in German graphology.

The matching method employs strict controls, is entirely objective in scoring, and does not limit the graphologist's freedom of expression. The technique has, however, certain shortcomings which may be briefly listed:

(1) If a judge's first matching is incorrect, he automatically makes two mistakes. Each further error makes

his other judgments less likely to be correct. Thus the order in which the sketches are read is likely to affect the result; a good likeness might be recognized if it came first, but, owing to previous mistakes, might fail to be recognized if it came late in the series.

(2) Many of the correct and incorrect matchings are, no doubt, based on a process of elimination rather than on the apparent resemblance of the graphological sketch and the personality. Both these criticisms can be overcome, however, if we provide, say, twice as many sketches as there are graphological specimens.

(3) Vernon has shown in a different connection (*187*) that the differences in ability of the judges to make correct matchings is considerable. Therefore, to average the results of several judges does not represent the method at its best.

(4) It is not easy to eliminate all the spurious elements which make for correct matching. For instance, mention of the racial ancestry of the writer, his address, or certain facts that are known to both the graphologist and the judge, produces an illegitimate basis for identification.

(5) The next chapter will show that "good" and "bad" errors are possible, and are not readily allowed for in available methods of scoring the success of matchings. For instance, two highly similar personalities may be confused with one another by the judges; such an error ought to weigh less than the confusion of two very unlike persons.

(6) The optimum number of subjects in each group has yet to be determined. It would also be valuable to know the effect on the statistical probability of the results if second or third choices were allowed. Tables for the probability of various types and combinations of matching should be prepared by psychological statisti-

cians, for the method has utility in psychological research beyond the field of handwriting.

All in all, matching must be regarded as a fruitful technique. The next two chapters will describe original experiments in which it is employed, and the results though less striking than those obtained by Bobertag, abundantly justify its use.

The Psychology of Handwriting and Its Bearing on Graphology. Graphology is the art (and perhaps embryo science) of determining qualities of personality from script; the psychology of handwriting, on the other hand, has no special interest in the diagnostic use of handwriting, but concerns itself merely with analytical and psychological studies of graphic movement. Much of its work, however, yields results of practical significance for graphology, as well as for law, education, and psychiatry; above all the psychology of handwriting contributes to a general understanding of the fundamental problems of expressive movement.

Chapter V contained a summary of results obtained from studies of the pressure of handwriting, and at the end of Chapter VI a demonstration was given of the value of this approach to graphology. The reader will recall that inexperienced judges achieved marked success in matching pressure-curves and handwriting specimens with brief sketches of personality. Inquiries among these judges showed that the more successful ones based their judgments on the records of pressure rather than on the scripts. This simple experiment, then, indicates that a knowledge of the psychological and physiological processes that occur in handwriting is perhaps even more helpful than a study of the shape of the writing itself.

It was pointed out in Chapter I that Downey's comparison of graphic characteristics with gestures, walk, and

carriage, is probably the only previous study in America of the consistency of expressive movement *per se*. Her technique was described on pp. 14–16. Eleven judges rated the motor characteristics of certain subjects from observing their gait and gesture. Downey rated the same characteristics from a study of the subjects' handwriting. Only two alternatives were allowed in the rating (rapid-slow, impulsive-deliberate, etc.). She found that the ratings on gait and gesture agreed with the ratings on script in 57% of the cases, and disagreed in 43%. The two point scale, however, is too crude to secure the maximum agreement which might be obtained by the method of independent ratings. Other experimental difficulties in this work were discussed in Chapter I, and it was shown that with slight changes in technique higher agreements might result.

In another experiment Downey (*39*) gives evidence that mood has an appreciable influence on script. She and one other subject wrote their signatures every day for four months, together with a brief account of their mood and state of health. At the end of the period, Downey measured various features of these scripts, and found that an increase of "total graphic movement" accompanied energetic moods, while slant and alignment were irregularly affected by changes in emotional condition.

Another important study of slant has been made by McAllister (*120*) showing that more time is required for writing centripetal pen strokes and a backhand slant. The significance of such findings is taken into account by Saudek, who employs slant as an indication of emotional traits *after* due consideration of the style of penmanship which the subject learned. Here, for example, we have an illustration of the way in which analytical

studies of movement contribute to the practice of graphology.

Studies in the psychology of handwriting underlie Downey's well-known Will-Temperament tests (*40*). The literature bearing on these tests is too great to be included here; Uhrbrock (*184*) summarizes it up to 1928. The W-T tests measure chiefly the speed and modifiability of writing under various conditions, but not the shape, style, or detailed features of the script. It is the impulsiveness, flexibility, and restraint of activity that are sought in this test. Marked individual differences are found, and the constancy of the subject's expressive activities is shown by the fairly satisfactory reliabilities; but the validity of these tests as methods of determining traits of personality is not fully established.

Hennings' unique *Zweipersonen Experimente* (*67*) make somewhat analogous use of writing for a study of such characteristics as flexibility, suggestibility, and imitativeness. One of his methods is to have two children write simultaneously, the script of one being reflected through prisms, or epidiascopically, onto the paper of the other. He then observes the extent to which one child's writing, intentionally or unintentionally, takes on the characteristics of the other's.

The consistency of expression in different spheres of motor activity may be further illustrated by comparisons between script and speech. Scripture (*162*) has shown that in normal speech a suitable amount of energy is applied for the conveyance of words to the listener, so that excessive energy of articulation indicates great vitality in the organism as a whole, either as a permanent trait,[1] or as temporary emotional excitement.

[1] This interpretation is confirmed by the high correlation reported in Chapter V, between ratings on Voice Intensity and all other tests and ratings which depend on "general vitality."

This idea concerning the discharge of energy in vocal expression fits in with the theory of temperament on which the Downey W-T tests are based. Wagoner and Downey (*191*), therefore, developed a set of speech W-T tests, measuring the speed, modifiability, and other characteristics of speaking. They found some high, some low agreements with the corresponding writing tests. For instance both motor impulsion and inhibition correlated +.63 in the two spheres of expressive movement, but the other measures gave lower agreements. In a further study, Wagoner (*189*) showed that W-T scores for graphic movement are characteristically different for groups of stutterers, college debaters, people who drawl or hesitate in talking, careless and negligent speakers, and precise speakers.

Saudek (*148*) also draws many parallels between general features of writing, speech, and thought; *e.g.*, "we write indistinctly and confusedly when we think indistinctly and confusedly," and there is "an inseparable parallelism between language and writing." Precise and pregnant brevity of speech goes with restrained, upright writing, verbosity with frank, broad, slanting hands. In Germany, Román-Goldzieher (*143*) describes the unique features in the writing of stutterers; their script like their speech tends to be arhythmic and "dismembered."

The importance of handwriting for certain legal and pedagogical purposes is unquestioned. On the determination of the authorship of questioned documents Osborn's book (*126*) is the standard treatise. In the educational field, we need only mention the writing scales of Ayres, Thorndike, and the varied studies of Freeman. For our purposes all this work is significant chiefly for its elevation of graphic movements to a position of dignity among the fields of psychological research.

Quality of writing, as measured by standard scales, or as judged by teachers, has been compared with intelligence test scores by several investigators. Gesell (*58*) and Starch (*170*) found correlations of about +.31 among children. But Thorndike (*181*), working with adults, obtained zero agreement. In Omwake's study (*125*) quality of writing according to Thorndike's scale was compared with Army Alpha scores, the correlation being +.047; and neatness, as rated by five judges, gave a similar low figure. The excellence of chirography would seem, then, to depend on intellectual ability only in immature subjects, if at all.

Experiments on hereditary features in writing seem to confirm the hypothesis that writing is a product of central determination. Siblings, for example, show objective similarities in their handwritings (Starch *171*), and these similarities are still more marked among twins (Kramer and Lauterbach, *91*). Studies by Gesell (*59*) and by Downey (*37*) confirm the opinion that handwritings in families have more than merely adventitious resemblance. All these findings strengthen the *a priori* case for graphology.

Limitations of Graphology. This broad survey has disclosed numerous arguments, both logical and experimental, for considering graphology a legitimate and promising field for psychological exploration. (This conclusion has no necessary relation to the question of the commercial value of graphology.) At the same time, there are certain fundamental questions bearing upon the ultimate limitations of graphology which should be explicitly stated. As yet no one knows whether these limitations are great enough to disqualify graphology for practical use, but they must be carefully considered by those who are interested in its theory.

It is, for example, a question whether *all* features of personality should be expected to express themselves equally clearly in script. May it not be that certain traits, interests, and abilities can be more clearly determined than others? For example, insight, sociability, and religious faith, seem to be among the qualities which elude graphologists; whereas expansive, graceful, neat, emphatic, clumsy characteristics often seem to be deduced from handwriting with considerable accuracy. Is it the traits which have clearly marked motor concomitants that are most readily judged? (The reader will recall Saudek's skillful analyses of the four cases presented in Chapter VI.)

The general cultural or educational level of the writer is also one of the features which seem to be most plainly expressed in script. Crépieux-Jamin (*32*) has described the "signs" of superior and inferior mentality, and Saudek always looks for the "standard class" of the script before he proceeds to further analysis. The class, like speed, modifies the significance of every other indicator. His main criteria for judging class are (1) good aesthetic spacing of the lines on the page and of words on the line, (2) naturalness and spontaneity of the writing (artificiality being indicated by several indicators akin to those of slowness), (3) artistic and original deviations of the letter forms from those taught in the system of penmanship which the writer employs. Saudek classifies scripts into ten standard classes and maintains that no graphologist ought to work with scripts which differ markedly from his own class. He considers himself capable of interpreting the top six classes, but makes no apology for neglecting the lowest four. Similarly, he claims that "true psychology . . . will always be restricted by the subjective limits of the psychologist."

Apart from this meager evidence concerning cultural level and traits with motor concomitants, there seems to be little that can be said regarding the ability of graphologists to determine some qualities of personality better than others. It seems quite probable that there are distinct limits to the revelations which script can make even to the trained eye; but the nature of these limitations is still not fully understood.

According to present evidence there is another inherent limitation to graphological judgment: the graphologist seldom seems able to put his emphasis in the right place. One analyst, for example, writes that a certain subject would be a good "diplomat, surgeon, psychologist, mathematician, or painter." It is true that the subject in question is a cultivated young man, and that all these professions require cultural qualifications; but his interests (and interests are surely an important part of personality) are completely obscured in this analysis. As a matter of fact one of these interests is of major importance in the subject's life, but the graphologist cannot tell which one! An analyst sometimes refers to a writer's "nervousness," but whether this quality is a pronounced and outstanding trait or one that is minor and disputable, the graphologist is usually unable to say. In general it seems that personal characteristics which are reflected in handwriting are not reflected in proportion to their importance in the subject's life, or else, it takes superhuman skill to read the accents aright.

The validation of graphological diagnoses presents another problem. Some graphologists cannot understand why the psychologist is not convinced by testimonials from satisfied clients. Their decades of experience and their regular employment by large business firms seem to them fully adequate proof of their skill.

The average graphological analysis is especially difficult to validate owing to its unscientific and unstandardized terminology; equivocal wording is readily accepted by a credulous person, but imposes serious obstacles for the scientist. The classification of mental processes used by graphologists is distressingly unpsychological, for instance, their anachronistic "faculties" and their sharp distinction between the "logical" and "intuitive" types of mentality. Verbal self-contradictions appear frequently in the analyses; and the terms employed often seem to obscure rather than reveal the personality. In some of our investigations one-half to two-thirds of the points made by professional graphologists had to be discarded as incapable of objective confirmation or denial.

Against these criticisms graphologists urge that handwriting yields indications beyond the competence of experimental psychology. Just as psychiatrists detect trends in personality not accessible to the standardized approach of tests and ratings, so, they say, graphologists work with subtle patterns which elude objective detection. They hold that many of the indications in handwriting can only be validated in the future when the subject has expressed the trends now latent in his nature and disclosed for the present only to the graphological expert. Although this answer smacks of casuistry, no impartial psychologist will deny that before passing final judgment upon graphology he must make certain that the criteria employed for the purpose of validation are well adapted and reasonably adequate.

Conclusions. We see then that the ambiguous position of graphology in the world of psychology is due to the failure of both the graphologist and the psychologist to grasp the point of view of the other. Each has reason to regard the other as prejudiced and ignorant of his own

methods and aims. Commercial abuses and unscientific standards are, no doubt, prevalent. Graphologists are often unwilling to submit to adequate controls, though friendly coöperation, as Chapters X and XI of this book will demonstrate, is by no means impossible. What hesitation they show is easily excused when we consider the psychologist's inability to provide them with adequate material and facilities, or to devise methods which allow them proper scope for their "intuitive" procedure, or to secure an adequate criterion against which to validate their judgments.

It is clearly unreasonable for psychologists to take fright at the very mention of research in graphology. Indeed, with the exception of a few quixotic experiments, conceived by psychologists bent on rescuing the public from a somewhat imaginary ogre, most of the researches reported have yielded promising results. Satisfactory methods for analyzing graphic movements in the finest detail have been worked out; pressure, speed, and quality are studied objectively; information is available regarding the influences of heredity, temperament, training, age, and sex. The agreement of graphic movement with gait, gesture, and speech is studied, as well as the cross-transference of graphic habits, with results which tend to support the theory that motor activities within personalities are interrelated. Concerning the validity of graphological analyses of personality, there is evidence that controls which are stringent and at the same time permit reasonable exercise of the graphologist's talents, may eventually settle the problem to the scientist's satisfaction without necessarily overwhelming the graphologist with disaster.

CHAPTER X

MATCHING SKETCHES OF PERSONALITY WITH SCRIPT [1]

By *Edwin Powers*

Dartmouth College

The purpose of the following experiment was not to test the claim that certain specific graphological "signs" or formations have precise correlates in the single features of personality, but rather to determine to what extent the *form-quality*, the total impression made by the script as a whole, reveals *patterns* of personality. The method employed was the matching of specimens of handwriting with case studies which described briefly the personalities of the subjects. Though a number of graphologists were included among the judges, they were not called on to make analyses of the scripts, but only to try to identify the personality sketch which corresponded with each script.

The Subjects or Writers. The subjects were ten male adults, ranging in age from 21 to 45 years, the average age being approximately 32. Of these, three were graduate students, two were members of a university faculty, two were college graduates now engaged in business, one was a former college student of good standing who had just left college because of emotional difficulties, one was a postmaster, and one a barber. Two were foreign born, one a Korean and one Italian, but all had received the entire or greater part of their education in this country.

[1] This chapter is condensed from an investigation entitled *Graphic Factors in Relation to Personality*, 1930, Dartmouth College Library, Hanover, N. H.

Some of the graphologists quite rightly criticized the homogeneity of the group of subjects. All but two of the subjects had received college training in America, and the majority came from a similar cultural background. That the ease of matching is a direct function of the heterogeneity of the subjects will be demonstrated when the results are given below.

The Handwriting Specimens. Each subject copied a passage of 40 words on uniform paper, with his usual pen and in his natural manner. The passage was selected to include a number of letters such as *t*'s, *d*'s, *I*'s, and punctuation marks, whose formations are often alleged to be of characterological significance. No signatures were included. The accompanying cut (Figure IV) is a reproduc-

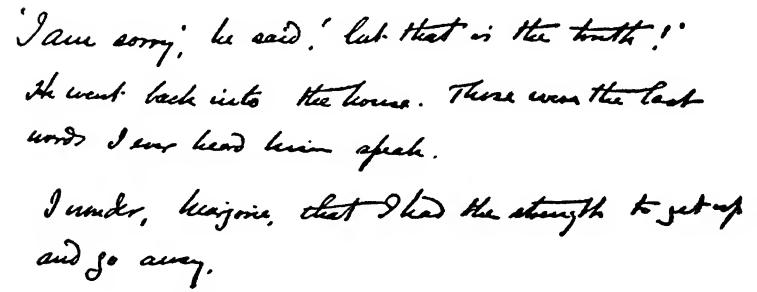


FIGURE IV. ONE OF THE TEN SCRIPTS EMPLOYED (two-thirds original size).

tion of one of the specimens; the appropriate personality sketch is also given. The subjects were unaware of the use to which their writing was to be put, and though it is not possible to know whether complete naturalness was achieved, yet the instructions given to the subjects sought to secure spontaneous scripts. Photostat reproductions were then made of each specimen, the natural size being retained. The right- and left-hand margins were unaltered;

while when the top or bottom margins were cut down (to economize space), the distance of each specimen below the original upper edge of the paper was indicated on the reproductions, so that the nature of the spacing adopted by the writer was always known.

SKETCH OF THE PERSONALITY OF THE WRITER OF THIS SCRIPT

A———though only 31 years old, is already one of the leading authorities in the world on certain aspects of early theological doctrine. He is an expert on manuscripts in several languages. He holds advanced degrees from Harvard and Oxford. His mother comes of an old family, and his home is filled with early American antiques and family memories. In such an environment he grew up, gradually increasing in contempt for his fellows who continually made life miserable for him at school and in the streets. He is timid physically. He cannot drive a horse or an automobile very well, chiefly because he is afraid of getting into difficult situations or getting lost. He exaggerates slight illnesses and consumes much medicine. When he finished theological training he repudiated ordination, and put his knowledge to use by writing on the disputes of the ancient Church Fathers. He went to England, immediately fell in love with English ways, and never in England has he been taken for an American. He is thoroughly English in manner and prefers to live there. He works speedily and effectively, but is restless. He is totally incompetent along mechanical lines, and his one accomplishment in sport is his brilliant game of chess. He ill conceals his boredom at others' comparative slowness unless he puts himself out to be charming, when he succeeds very well. He has quick appraisal of objects of beauty, the worth of ideas, and the nature of people. He discloses a buried sensitiveness when he talks about himself, which is seldom. Before a friend can speak words of sympathy and understanding he has set his problems aside with a witty remark. His intellect, which is sharp and satirical, prevents the needs of his nature gaining too strong a hold over him. When asked to give a sample of his handwriting for the experiment he wrote: "I enclose a specimen of my handwriting and hope devoutly that you will be wholly unable to deduce anything about my character from it. If I expose the horrid truth every time I take my pen in hand, I must resort to a Remington even for signing checks."

Of 13 graphologists who later commented on the experiment, 7 objected to the fact that the script was a copy, 4 regretted the lack of signatures, 4 considered the length of the script insufficient, and several claimed that the use of photostat copies instead of the originals added to their difficulties. If the copies were unsatisfactory, however, the graphologists were informed that the original script would be mailed on request for a more careful consideration. Only two graphologists made such request.

The Sketches. A sketch of about 250 words was made of each subject by three psychologists in collaboration, at least one of whom was intimately acquainted with the subject, the other two having had some personal contact with him. The writers of the sketches aimed to delineate the outstanding traits as comprehensively and as accurately as possible, with proper emphasis on the subject's interests, values, attainments, and habits, in such a manner that the total or integrated personality would stand out clearly and would be readily recognizable to those who knew him.

Inasmuch as many professional graphologists claim to diagnose the "real self," the "true subconscious personality," which is hidden from the gaze of the external psychological observer, it is only natural that they should object to these sketches serving as the validatory criterion of the experiment. Nine of the 13 graphologists who commented on the experiment considered the portraits unsatisfactory for one reason or another, and it was found that those sketches most often objected to were on the whole those most often incorrectly matched with the script. But the same criticism applies equally to any criterion which the psychologist can offer. If the graphologist will not accept the judgment of competent observers

and friends, he will be even less likely to admit the validity of the findings of objective psychological tests, so that these objections merely lead to an impasse which would render hopeless the establishment of any kind of scientific graphology. While the writers of these sketches do not claim to know the subjects completely, and the experimenter admits that the sketches were of uneven merit, yet, on the whole, the sketches seem to be the least unsatisfactory criterion that can at present be devised.

The Experiment. The specimens of script were numbered from 1 to 10, the sketches lettered from A to J, in a different order. They were then submitted to the judges with the request that they identify the writers by matching each specimen with its appropriate sketch. In this way complete objectivity of scoring was obtained.

The following judges were enlisted for the experiment:

(a) One hundred forty-three male undergraduates, most of them in their sophomore year. All but three of these denied that they had any previous training or experience in graphology, the three claiming only a passing familiarity with the topic.

(b) Twenty-five members of the college faculty, or their wives, representing eight college departments. Of these all but one denied that they had ever made any study of graphology, and all but two disclaimed any ability as amateur graphologists.

(c) Seventeen professional graphologists, eleven of them women, including some of the more prominent professionals in this country and two in Europe. Several of them had been engaged in this work for more than thirty years.

The Number of Correct Matchings. The number of correct matchings is given in the following table for each of the three groups of judges, and for all 185 judges combined. The first column represents the number correct out of 10, while in the succeeding columns are given the numbers and the percentages of judges in each group achieving these scores.

NUMBER AND PERCENTAGE OF CORRECT MATCHINGS

CORRECT MATCHINGS	UNDERGRADUATES $n = 143$		FACULTY $n = 25$		GRAPHOLOGISTS $n = 17$		COMBINED $n = 185$		CHANCE $n = 185$	
	No.	%	No.	%	No.	%	No.	%	No.	%
0	24	16.8	4	16	1	5.9	29	15.7	68	36.8
1	44	30.8	7	28	2	11.8	53	28.6	68	36.8
2	34	23.8	7	28	7	41.1	48	26.0	34	18.4
3	28	19.6	5	20	4	23.5	37	20.0	11	6.1
4	10	7.0	1	4	2	11.8	13	7.0	3	1.5
5	1	0.7	1	4	1	5.9	3	1.6	0	0.3
6	2	1.4	0	0	0	0	2	1.1	0	0.1
7–10	0	0	0	0	0	0	0	0	0	0.0

The mean scores (numbers of correct matchings) are 1.77, 1.80, and 2.41 for the three groups; and 1.83 for all the judges combined. In other words 339 (18.3%) of the total 1850 trials were correct. According to the theory of probability, an average of 1.0 correct matchings could be obtained by chance, so that each average is somewhat superior to chance. In order to evaluate more exactly the statistical significance of the results, a computation was made under conditions similar to those of the experiment.[1] The following table should be read: 0 or 1 correct matchings would occur by chance once in 2.7 times; 2 correct matchings would occur once in 5.4 times, and so on.

[1] By Professor B. H. Brown of the Department of Mathematics of Dartmouth College.

Number of Correct Matchings	Chance Frequency of Occurrence
0	2.7
1	2.7
2	5.4
3	16
4	65
5	327
6	1,920
7	15,120
8	80,640
9 and 10	3,628,800

It will be seen that the probability of getting 1.83 correct matchings (the average score) by chance lies at about 1 in 4.5 times. For the group of graphologists the probability of getting 2.41 correct by chance is about 1 in 8.5. The chance of getting 6 correct, the score obtained by two of the students, is 1 in 1920 times.[1]

The Frequency of Correct Matchings for Each Pair. Another method of expressing the results is to list the number of times each pair (script and sketch) was matched correctly by the judges. For the ten pairs, the proportions of correct matchings are given below; the subjects, with their code letters and numbers, are listed in the left-hand column.

It will be seen that the range of correct matchings per subject is from 58.8% for F-3 (made by 10 out of the 17 graphologists) to 0% for C-7 (made by the same group). By chance expectancy, 10% would be correct, so that combined groups show a pairing above that of chance for 7 out of the 10 subjects.

It should be noted that the correct pairing most frequently made by one group was not that most frequently

[1] The statistical significance of these results is stated by an alternative method on pp. 229 f., in conjunction with the results of a later experiment which in all essential respects is confirmatory.

made by either of the other groups. The rank order correlations between the apparent "order of difficulty" of the 10 pairings are: undergraduate-faculty +.88; undergraduate-graphologists +.61; faculty-graphologists +.74. The undergraduates, therefore, seem to resemble

FREQUENCY OF CORRECT MATCHINGS

SUBJECT	UNDER-GRADUATES	FACULTY	GRAPHOL-OGISTS	COMBINED
Barber (E-1)	47.6%	36%	52.9%	46.5%
Faculty, theological scholar (A-10)	21.0	48	41.1	26.5
Graduate student Korean (G-6)	23.1	20	35.3	23.8
College student (maladjusted) (F-3)	16.1	16	58.8	20.0
Graduate in commerce (H-4)	18.9	24	11.7	18.9
Graduate student (C-7)	16.1	8	0.0	13.5
Postmaster (I-5)	11.9	12	17.6	12.4
Faculty, psychology instructor (J-8)	9.8	4	5.9	8.6
Graduate student (D-2)	8.4	8	11.7	8.6
Graduate in business (B-9)	4.1	4	5.9	4.3

the faculty closely in this respect, while the graphologists are relatively unlike both the untrained groups.

When the general distributions of the matchings are tabulated, *i.e.*, the number of times each possible combination of sketches and handwritings are chosen by the three groups of judges, further interesting facts emerge. Among the undergraduates the "modal" judgment is correct for 4 subjects; that is to say, of the 10 sketches most frequently assigned to the 10 scripts by the student

judges, 4 were correct. Among the faculty and graphologists, the corresponding figures are 3 and 4 out of 10. But when all the groups are combined, the modal pairing is correct 5 times out of 10. This is evidence of the greater correctness of combined judgments over the average of the judges' separate scores.

Qualitative Results. An examination of the tables seems to show that the general distribution of judgments does not at all resemble what we might have expected from a chance selection of pairings. The results could not reasonably have followed if there existed no individuality or form-quality in the writings. For example, the subject whose sketch and script were most frequently matched was the barber; he was described in the sketch as considerably below the intellectual and cultural level of the other subjects. The data show, moreover, that while his handwriting was paired correctly with his sketch in 46.5% of the judgments, it was paired with the sketch of the brilliant theological scholar in only 2.2% of the judgments. It is interesting that the faculty group achieved its highest score in the case of the same scholar. There seems to be a certain form-quality about the handwriting of a scholar which educated people more readily perceive. The theologian was mistaken for the postmaster ("an efficient politician . . . an unimaginative thinker") by only 3 judges, and the postmaster's script was matched with the scholar's sketch by only 4.

A similar example was the handwriting of No. 4, the largest, perhaps the most expansive and striking. This was correctly matched with the sketch of "one of the leading representatives of the U. S. Department of Commerce abroad" 18.9% of the time. Only one judge (less than 1%) believed it to be the handwriting of subject G, "shy, reserved, silent Oriental." Only two judges con-

sidered that it belonged to subject C, "graduate student in his early twenties . . . even in temper," and only three identified it with E, "the owner of a modest, orderly three-chair barber-shop in a small town." On the other hand, 47 or 25.4% of the judges, including 10 of the 17 graphologists, picked out this handwriting as that of subject B, whose "very striking personality centres around a strong emotional life . . . often melodramatic . . . extremely frank, uninhibited and sympathetic." While an almost equal number (45) thought that it must belong to D, a man of "very strong and large physique, excellent health, and enormous vitality."

Numerous other cases could be cited from the tables which suggest that individual differences in handwriting are judged to represent personality differences in a consistent fashion. Where any large errors occur, they seem to be "good" errors: the more unlike two personalities, the less frequently are their handwritings mistaken, and the more similar they are in some essential characteristic, the more often are their handwritings confused.

Attitude of the Graphologists. In spite of the reputed aversion of professional graphologists to scientific control, the experimenter found them on the whole sympathetic toward this investigation. Only 4 out of the 21 to whom he applied refused to take part in an investigation which they knew would deny them their usual freedom of expression. Practically all the objections mentioned above were made before they knew their own scores. Many expressed no confidence in their matchings owing to the conditions of the experiment, while 13 of the 17 were sufficiently interested to give specific reasons why they did not consider it an adequate test of their graphological ability, in several cases carrying on correspondence relating to the experiment for a considerable period of time.

The general attitude of many of the graphologists is expressed in the words of one of them, as follows, "Quite frankly I don't think these experiments are a true test of the worth of graphology, but I am very much interested and shall await the results with an open mind."

Summary. 1. The ability of untrained persons to match specimens of handwriting with personality sketches of the writers, as measured by this experiment, is somewhat above what one would expect by pure chance. The mean number of correct matchings out of a possible 10 was found to be 1.77 for the student group, and 1.80 for the faculty group; according to probability we would expect a mean of 1.0.

2. The skill of a trained group of professional graphologists in the same test exceeds the untrained group about as much as the latter exceeds chance. The mean for the graphologists was 2.41 out of 10.

3. The greatest number of correct matchings by any judge was 6, made by two of the student group. One member of each group made 5 correct matchings. How much better these scores were than chance we can gauge by referring to probability tables, which tell us that the chance of any one judge matching only 1 correctly is 1 in 2.7 times; two correctly, 1 in 5.4 times; six correctly, 1 in 1920 times. A perfect score will be made by chance once in 3,628,800 times.

4. Combined or modal judgments are more correct than the mean of the individual judgments. That is, taking the collective judgments, we find that in 5 out of the 10 choices, the judgments most frequently made were correct; in 7 out of 10 cases the modal scores exceeded chance.

5. The general distribution of matchings is unlike that of a chance distribution; the form-quality of handwriting

seems to influence the selections. Handwritings evidently indicate to the judges definite personality patterns, and though these are not always the correct ones, the errors for the most part appear to be logical and meaningful.

6. There are certain advantages in the application of the matching method to the investigation of graphology. The method does not confine the judge to the study of any of the single "signs," but gives him an opportunity to pass judgment upon the total sample of script. The problem of evaluating subjective graphological analyses is avoided, and in respect to scoring, the method is entirely objective and strictly quantitative.

7. On the other hand, certain limitations in our application of the method are apparent.[1] Some of these were swiftly pointed out by the professional graphologists. The inadequacy of the sketches, lack of signature, unnaturalness of the writing, brevity of copy, homogeneity of the subjects, and the use of photostat reproductions, were all mentioned by them as handicaps. Whether the low scores on the test were due to the limitations of the method or to the limitations of graphology cannot be decided dogmatically.

8. This experiment illustrates, contrary to current belief, a willing and coöperative attitude on the part of professional graphologists in a scientific investigation of their skill.

[1] For a critical discussion of the matching method see pp. 197–203.

CHAPTER XI

GRAPHOLOGICAL AND PSYCHOLOGICAL ANALYSES OF THE SAME PERSONALITIES

The following experiment is, in a sense, the reverse of the one described in the previous chapter. There the graphologists were required to match specimens of handwriting with the case studies written by psychologists; here the graphologists are allowed to write their own case studies in any way they wish from the evidence of the script, and the psychologist attempts to identify these studies, using his knowledge of the subjects gained through personal contacts and through a long-continued study of their performances on a variety of psychological tests. The experiment was made possible through the generous coöperation of Mr. De Witt B. Lucas, a well-known professional graphologist, and Professor June E. Downey, who has written extensively on the psychology of handwriting but does not claim any special skill in judging personality from script.

The Subjects or Writers. The subjects were 23 college freshmen, aged 17 to 19 years. They had been tested during 1930 for about 20 hours with most of the available personality tests; most of their test scores were found to possess considerable theoretical and empirical validity. In addition, these subjects had been rated by friends, by six faculty members, three psychological observers, and by two floormen of the University Commons where all of them worked. The experimenter supplemented these data by means of frequent association and interviews, and by subsequent correspondence. Draw-

ing upon these various sources, full personality sketches of the subjects were prepared, as well as profiles of their scores on the tests.

The completeness of the data available for each subject gave an excellent opportunity for this graphological experiment. The only serious shortcoming was the immaturity of the subjects; this difficulty was emphasized by both analysts in their discussions of the results. Offsetting this disadvantage to some extent was the fact that the subjects were as heterogeneous a selection as is likely to be found in college. They had been specifically chosen with reference to diversity of family income and cultural background, place of birth, fields of study in college, and other conditions.

The Script. Specimens of the subjects' handwriting were taken from their test blanks, or obtained at various times over a period of twelve months previous to the graphological analysis, without the subjects having any idea as to the purpose to which it was to be put.[1] The specimens consisted of portions of essays, answers to group intelligence tests, postcards written to the experimenter, etc., and included half a dozen or more signatures for each writer. The content of most of the script was identical for each writer, and it was written on identical varieties of paper; in very few cases could anything be deduced from style of composition; words or sentences which by their content might have thrown some light on the personalities, were blocked out. Since most of the material was obtained before the experiment was planned,

[1] For the purpose of obtaining spontaneous script the experimenter recommends the "memory test." The subjects were given a short story to read, and were later asked to write out all that they could remember. Since they believed that their ability to recall the story was being tested, their writing was probably more natural than it would have been in copying, or writing from dictation. And at the same time the content or literary style of the scripts cannot have given the graphologists any very definite clue to their personalities.

the subjects had been permitted to write in pencil; while the analysts did not find fault with the spontaneity of the script, both wished that more of it had been written in ink. The material included, for each subject, a blank on which he had recorded his age, address, schools attended, and his right- or left-handedness. Mr. Lucas received a page or two of ink-written script which was not available to Professor Downey; and she alone received the subjects' test blanks, time records, and other data, for several of the Downey Individual Will-Temperament tests.

The analysts were given complete freedom to work in any way they desired, and in writing their personality sketches to use any terminology they wished. They were asked, however, not to mention the racial ancestry, address, or any other objective evidence obtainable from the data which might aid the identification artificially. They both numbered their analyses according to a key unknown to the experimenter, so that his only clue to the identity of each analysis was through its agreement with his personal knowledge of the subjects, and its resemblance to his own, previously written, psychological case studies.

In the subsequent discussion, the two sets of 23 analyses or sketches will be referred to as the L analyses and the D analyses (*i.e.*, Lucas' and Downey's respectively). The L and D experiments were conducted independently a month or more apart; but almost identical techniques were used in validating both sets of analyses, so that the results may be presented together.[1]

Identification of the Entire Group of Analyses. The experimenter first tried to identify all the 23 L analyses,

[1] In order that the reader may form a concrete picture of the material upon which this experiment is based, two typical case studies for one subject (S. I. G.) are presented. According to the criteria employed by the experimenter both the L and D sketches, although of unequal length, are about equally correct, and

making a first choice, usually a second choice, and occasionally a third choice; thus he assigned an average of two names to each sketch. Profiting by the experience gained in this experiment, he made half a dozen choices for each of the D analyses, arranging them in order, so that the rank of his correct choices could be determined. The following results were obtained.[1]

the degree of correctness for each analyst is quite near his or her median for all the 23 sketches.

D's sketch: This is an individual of superior intelligence and much imagination. There is, however, a certain amount of timidity in his make-up and lack of self-confidence. He is very adaptable in his dealings with people and compensates for his failure in direct aggressiveness by subtlety in approach and indirect manipulation. He has an artistic bent of some sort, but is too timid to let himself go in an imaginative way.

L's sketch: He is fine grained, exceedingly well endowed and well developed for his age. His writing shows many fine inherent qualities that will make a strong character and win him a worth-while objective in life, if he gains a proper understanding of life and its responsibilities and exactions. His is a decidedly scientific mentality, scholarly and one which will digest and thoroughly assimilate his ideas. He is wonderfully awake, alert and conscious of what is transpiring around him, has a natural penchant for criticism and analysis, and would make an excellent chemist, surgeon, physician or special worker, along lines of scientific analysis and investigation. He is one who will profit from a college education. His judgment is inherently good, but objectively undeveloped. He is naturally economical with great facility of adaptation, but has not yet come into full possession and conscious use of these characteristics. He is not lacking in enough defensive finesse to protect himself and his interests, or the interests of others trusted to his keeping. He is a good instinctive reader of people and character, if he heeds his intuition, which is exceptionally represented for a man. He understands readily, even though he may keep back any outward show of comprehending at the time. He has a tendency to allow the other fellow to commit himself and to see the other fellow's cards, before laying down his own hand. A young man of great capacity, facility, inventiveness, versatility, and resourcefulness. These qualities are largely latent at present, but will be awakened in time and serve him well. He is not so selfish, as he is economical, careful, and adaptable. He could easily carry these qualities too far, and become penurious and miserly in time, so gradually that he would not realize it. There is a tendency to become discouraged as he proceeds in a difficult task, and to wind up depressed and lacking in the original enthusiasm and interest of his beginnings. He is not conceited, vain, shallow or pretentious. In some ways, he is over-modest in a self-depreciatory sense. Has general good health and normal vitality, that has not been weakened by vicious dissipation. Friendly and approachable, yet highly discriminating and exclusive in his attachments. Affectionate, yet well controlled in matters of the heart. His observation, imagination, constructiveness and intuition, will make him capable of performing unusual kinds of work, and even pioneering in new or strange occupations or avenues of investigation. Literature, medicine, psychology and science are recommended as occupations.

[1] Two of Professor Downey's analyses were omitted in this portion of the experiment, since they contained objective clues, and their correct identification was therefore illegitimate. In the validation of particular traits (discussed below) all 23 analyses were employed legitimately.

	23 L Analyses	21 D Analyses
Number of correct		
first choices	3	1
second choices	2	3
third choices	0	0
fourth choices	0	1
fifth choices	0	2
sixth choices	0	1
No correct identifications	18	13

It would be theoretically possible to find the exact probability of this degree of successful identification; but since different numbers of choices were assigned to many of the analyses, the method would be unnecessarily complex. The following approximation is more convenient and sufficiently accurate. Each correct second choice was considered as equivalent to one-half of a first choice, each third choice to one-third, and so on. This plan then gives us four correct matchings for the L analyses and 3.317 correct matchings for the D analyses.

Now it can be shown that in a problem of this type, where N is the size of the population to be matched, and n is the number of correct matchings, that P, the probability of obtaining one, two, or three, etc., correct matchings, is equal to:

$$\frac{1}{n!}\left\{\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}-\frac{1}{5!}+\cdots+\frac{(-1)^{N-n}}{(N-n)!}\right\}$$

When $(N - n)$ is greater than about 10, this equation reduces to,

$$P=\frac{e^{-1}}{n!}$$

The probability of getting 4 G and 3.317 D sketches correct is approximately .01533 and .042 respectively. In other words this degree of success might occur once in 65.2, and once in 23.8 times by chance.

Identifications of Smaller Groups of Analyses. The 23 L and the 21 D analyses were then divided for the experimenter at random into three groups each, and he was told the names of the subjects in each group, though not the correct order of their names. The experimenter then attempted identifications within each group; first, second, and occasional third choices were assigned to each sketch, with the following results.

	L Analyses	D Analyses
Number of correct		
first choices	5	5
second choices	2	3
third choices	0	3
No correct identifications	16	10

Giving second choices half weight, third choices one-third weight, as before, we find that the average number of correct first choices is 2 out of 7.67 for the L subgroups, and 2.5 out of 7 for the D subgroups of sketches. The probability of obtaining these results by chance is approximately .1833 and .110. These results might occur once in 5.45, and once in 9.09 times. This experiment is therefore rather less successful than the previous one.

The significance of these results may be more readily grasped when they are expressed in terms of their P.E.'s. Taking both Powers' investigation and the present experiment, the following table gives the values of P which have already been cited in Chapters X and XI; in the next column is $\frac{1}{2}(1 + \alpha)$; in the next column is given x, as determined from the Kelly-Wood or Sheppard tables; and in the last column is the number of times each result exceeds its P.E., *i.e.*, $\frac{x}{.6745}$.

We see that in each of the experiments the result is nearly two, or more than two times its P.E., but that a

THE PROBABILITY OF CORRECT MATCHING

INVESTIGATION	CORRECT MATCHINGS	P	$\frac{1}{2}(1+\alpha)$	x	TIMES P.E.
Powers: average result	1.83 out of 10	.208	.896	1.26	1.87
Powers: graphologists' result	2.41 out of 10	.140	.930	1.48	2.20
Powers: best result	6 out of 10	.00052	.99974	3.47	5.14
Lucas' sketches	4 out of 23	.01533	.99234	2.42	3.59
Downey's sketches	3.317 out of 21	.042	.979	2.04	3.03
Lucas' sketches	2 out of 7.67	.18333	.90833	1.33	1.98
Downey's sketches	2.5 out of 7	.110	.945	1.60	2.37

conventionally (*i.e.*, statistically) significant result is obtained only in the identification of Lucas' 23 sketches, and in the case of the most successful of Powers' subjects.

Identifications by the Subjects Themselves. Of the 23 subjects, 12 were selected who were known to be fairly well acquainted with one another. In this selection the most uncomplimentary L and D analyses were avoided; but no account was taken of their probable accuracy or inaccuracy. Separate copies of the two sets of sketches were renumbered and sent to the subjects; a list of the twelve names, in a different order, was given to them, and they were asked to try to identify, first the two analyses (one L and one D) which represented themselves, and then to pick out those analyses which fitted any of their friends. Second choices were also encouraged. Stress was laid in the directions on avoiding the influence of the complimentary or uncomplimentary nature of the analyses.

Eight subjects replied, sending in a varying number of judgments. Second choices are given half the weight of first choices in the following table. In parentheses are printed the total number of judgments submitted.

	L	D
Correct choices of own analysis	½(10½)	½(9½)
Correct choices of analyses of others	7(48½)	4½(43½)
Combined total	7½(59)	4½(53)

Thus only 12 out of 112 choices, or 10.7% were correct. By chance, however, one-twelfth, or 8.3%, might be correct, a figure which is nearly as large as that actually obtained. Applying the same technique as in the previous section, we find that the result is 1.5 times its P.E.

In spite of this poor result, it is probable that chance is not a sufficient explanation, since out of the 12 correct judgments, 5 were supplied by one subject, 3½ by a second, 2 by a third, 1 by a fourth, a half by a fifth, and none by the others. This would seem to show that some subjects can identify a few of their friends in these graphological analyses, while others are totally unable to do so, rather than that no subject is able to make judgments whose validity is much superior to chance.

One should note, in concluding this part of the experiment, that, as in Powers' study, there was a fair amount of agreement among the different judges in their choices, although these modal choices were not always correct. As is frequently the case in judging personality the same false cues are seized upon by many of the judges. The results point to systematic error rather than to random guessing.

Validation of Particular Traits. After the completion of the identification experiments, the graphologist picked out certain subjects who, he thought, would rank most highly in the group with respect to certain "traits" of his own choice. In this way he selected:

the four most intelligent, capable and generally resourceful subjects;
the three least reliable subjects;
the three subjects with the worst health;
the two most inconstant and unstable;

the four who showed the greatest emotional intensity, reacting mainly through their emotions;
the four with the greatest "head direction and mind rule";
the four most sensitively reserved and introverted;
the four who show the greatest tendency to extroversion;
the one subject who is the best balanced in every respect.

Further descriptions of these traits were given by the graphologist. It will be seen that he made 29 selections in all. Before the experimenter saw the selections, he attempted to rank all the subjects according to this list of traits. As far as possible he based his rankings on the objective scores for intelligence, extroversion-introversion, emotional instability, and the like. A comparison was then made for each trait between the ranks given to the subjects by the graphologist and the ranks given to the same subjects by the experimenter.

We may assume that if there was only chance agreement between the graphologist's selections and the experimenter's judgments, then the average rank of all the graphologist's choices would be median on the experimenter's list (12th in the series of 23 cases). In the following table the correctness of the graphologist's selections for each trait is expressed as a percentage; if his selections for a given trait corresponded to the experimenter's, that trait would receive a score of +100%, and if his selections were the reverse of the experimenter's the score would be −100%.

Inconstancy, instability	+70%
Reliability	+68
Introversion	+25
Ruled by his mind	+17
Intelligence	+5½
Emotional intensity	+ 3
Extroversion	−31
Poor health	−51
Best balanced	−62

The average success of all his selections is +7.3%. The statistical significance of this result was determined as follows. The difference between the average ranks of the 29 subjects whom the graphologist selected, and all the others whom he did not select on these nine traits, was 0.862 ranks. But the standard deviations of the ranks of his selections were very large, so that the P.E. of the difference is .89. The difference is therefore barely equal to its P.E., and could easily occur on the basis of chance.

The result though negative should not be considered conclusive. In the first place, the "traits" as named by the graphologist and understood by the psychologist may be quite different. Furthermore, the experimenter's criterion for several of the traits was largely subjective, and there is no means of telling whether he or the graphologist was nearer to the truth. Finally, it should be noted that selections on some traits are far better, according to the criterion, than on others; though it is unsafe to draw the conclusion on the basis of these few cases, that the graphologist and the psychologist are always likely to agree well on, say, "instability," or always disagree on "best balanced."

A more objective technique, which yields results of fair statistical significance, was applied to the D analyses. Professor Downey was sent a schedule of the actual tests and scales, by which the subjects had previously been studied. She then selected (on the basis of her graphological analyses) the four, five, or more subjects whom she expected to stand at the top or bottom of the rank orders for several of these measures. Here then we not only eliminate the subjective criterion of the experimenter's judgments, but also all confusion as to the meaning of the traits involved. For example, Professor Downey did not select the "aggressive" subjects, her choices then being

compared with the experimenter's impressions, but instead she selected those subjects who, she thought, would be most likely to score high or low on the Allport test of Ascendance-Submission, and the experimenter already knew their actual scores.

There follows a list of some of the tests and scales used, together with the average rank of D's choices. If her selections had been determined merely by chance, they should all average 12th in rank. The percentages in the final column express the deviation of her judgments from this chance rank.

Introversion scores on the Neymann-Kohlstedt test	4.1	+83%
Submission scores on the A-S test	7.74	+50
Low intelligence (several tests and ratings)	8.0	+44
Maturity or organization of character (tests and ratings)	9.3	+30
Aesthetic or artistic interest and values (an interest blank, and The Study of Values)	9.6	+27
Average scholastic grades	9.63	+25
Conservatism (questionnaire test)	10.4	+18
Literary interests	10.6	+16
Economic or utilitarian interests and values	10.7	+13
Extroversion (Neymann-Kohlstedt)	11.0	+11
Gregarious and social interests	11.2	+ 9
Low emotionality (several tests and ratings)	11.25	+ 8
High emotionality	12.25	− 3
High score on Meier-Seashore Art Judgment test	12.4	− 4
Ascendance scores (A-S)	12.4	− 6
Humanitarian interests and social values	13.0	−10
Low scores on Meier-Seashore	13.1	−12
High intelligence	13.8	−20
Radicalism (opposite of conservatism)	14.2	−24
Immature, unorganized character	14.2	−24
Political or power-seeking interests and values	16.2	−43

In addition to this series of judgments D picked out the least healthy member of the group correctly, but sug-

gested incorrectly the type of disease from which he suffers. Her three other choices for poor health showed no agreement with the rather unsatisfactory criterion of doctor's ratings.

The average of all these selections is +9.75%. Applying the method used in validating the L selections, it is found that the difference between the average rank of those whom D chose and the remaining subjects whom she did not choose on these 21 measures, is 1.1014 ranks, which is 2.23 times its P.E. This difference might occur 134 times in 1000 by chance, so that, though small, it probably does possess some significance. It should be noted that D's selections for some qualities are far more successful than for others. In making her judgments, she added an approximate rating of her degree of confidence; no differentiation in success, however, was found between her more and less confident choices.

Analysis of the Graphological Sketches. The two sets of sketches were next considered in detail; each separate point made by each analyst was isolated and marked +1, +½, 0, −½, or −1, according to its degree of agreement with the experimenter's knowledge of the subjects. The criterion here is, of course, unavoidably subjective; but as far as possible, only those points were marked on which the experimenter possessed some fairly definite objective evidence. Many points were either too vague to be verifiable, or dealt with matters upon which no evidence was available. This category includes 52.7% of the points in the L sketches. Since D employed more standard psychological terminology only 28.6% of her points were unverifiable by our data. In the following discussion and tables these unvalidated points are omitted. The actual percentages of the correct and incorrect verifiable points were as follows:

	L	D
Definitely right (+1)	11.0%	20.0%
More right than wrong (+½)	35.7	32.0
Equally right or wrong (0)	20.3	18.6
More wrong than right (−½)	26.9	18.1
Definitely wrong (−1)	6.1	11.3

Averaging the scores, the excess of right over wrong is 9.3% and 15.7% respectively.

Each subject's points were next scored separately. The range of scores was as follows:

	L	D
Most successful analysis	+43.8%	+60.0%
Median	+11.1	+12.5
Least successful analysis	−50.0	−35.7
Average of all analyses	+ 9.3	+15.7

In the two sets of sketches, 12 subjects in all received zero or negative scores, *i.e.*, an excess of wrong over right points; 34 received positive scores. No reason could be found why these 12 subjects should have been less correctly described (according to the criterion applied) than the rest; *e.g.*, they are not the youngest in the group, nor do they seem to show any other uniformity which might make them more difficult to analyze.

It was found possible to classify these verifiable points from both analyses into fairly definite and distinct psychological categories. In the following table the percentage of verifiable points in each category is given on the left, while the relative correctness, *i.e.*, the excess of right over wrong points is listed on the right.

Apparently there is but little similarity between the abilities of Lucas and Downey with respect to different categories. It was found that some of the subjects whom L analyzed most incorrectly were among those whom D analyzed best, and *vice versa*. The order of goodness of

Percentage of Points out of 322 L	Percentage of Points out of 150 D		Percentage Right Minus Percentage Wrong L	Percentage Right Minus Percentage Wrong D
6.5	7.8	*Physical characteristics:* health, liability to disease; size, physical type, quickness.	−11.9	+58.2
17.7	13.7	*Aptitudes and abilities:* including intelligence, "endowment," quickness of mental processes, mechanical and artistic talents, ability to judge personality, etc.	+10.5	+33.3
22.4	16.3	*Interests and values:* including professional interests, intellectuality, culture; spiritual, materialistic, athletic interests; also sentiments, prejudices, beliefs, ideals.	+18.0	+ 6.0
16.2	27.5	*Social traits:* gregariousness-reclusiveness, popularity, leadership traits, egotendencies, etc.	0.0	+14.3
19.2	18.3	*Emotional characteristics:* including stability, "temperament," moods, repressions, imagination and humor.	+15.3	+19.7
18.0	16.3	*"Character" traits:* persistence, dependability, industriousness, carefulness, efficiency, and their opposites.	+ 6.0	− 8.0

the L and D analyses of the same subjects correlated −.24. It is more than probable, however, that their methods of analysis were so entirely different that these discrepancies do not necessarily run counter to our statement (on p. 201) concerning the accord between different graphologists, as found by Schorn, Powers, and Bobertag.

Significance of the Results. The degree of success in this experiment for both analysts appears to be rather small, but the uniformity of the results is interesting. In Powers' experiment, we saw, the average matcher

achieved a success 1.87 × P.E.; in the identifications of the Lucas and Downey sketches, the figures were 3.59, 3.03, 1.98, 2.37, and 1.50 × P.E. In the selection of traits by Lucas and Downey, the results were 0.97 and 2.23 × P.E. The detailed analyses of the graphological sketches could not be treated as to their statistical significance, but they apparently yielded closely analogous results, which might well be from 2 to 3 times the P.E.'s, on the average. So that although any one of these results might seem discouraging, yet considered in aggregate they point to the conclusion that handwriting possesses an appreciable relation to personality.

An important qualification to all these averaged figures is the irregularity of success with respect to different subjects and different psychological traits. In both sets of analyses, very large differences, ranging from +70% or +80% to −50% or −60%, were found in the accuracy of judgments concerning different traits and subjects. Similarly, when the subjects themselves attempted to make identifications, some were far superior to others. Without further evidence, then, it would be unfair to state merely that most methods of validation prove graphological analyses to be only slightly superior to chance. On the contrary, a far more complex state of affairs exists, and at present we are quite unable to determine which psychological traits are most clearly expressed in script, what types of subjects are most accurately analyzed, or what graphological methods of analysis give the best results.

Several criticisms may be leveled against the techniques and material of the present study. The following defects may account in part for the rather unsatisfactory results.

(a) The subjects should be older, with greater graphic maturity.

(b) A larger selection of ink-written script should be placed at the disposal of the graphologists.

(c) The judges who attempt identification should be more numerous and should have known the subjects over a long period of time. Immature judges, such as the subjects in this investigation, who are only moderately well acquainted with the other subjects, are quite useless. In this study the experimenter's personal evaluations and judgments were his chief criterion. Another experimenter might easily have achieved a larger or smaller degree of success in identification, or might have scored the accuracy of the analyses somewhat differently. However, the similarity of his results to the results obtained by more objective methods tends to show that he was not markedly biased.

(d) In the matchings or identifications, it is probably better to use small groups of subjects, say 5 at a time; the judges should rank each subject according to his resemblance to each sketch, so that exact figures on the rank position of all the judgments may be obtained. This method would also throw light on the question of "good" errors raised in the previous chapter.

(e) The most objective, and unambiguous validatory technique is provided when an analyst, with both psychological and graphological training, scores the subjects according to their probable standing on standard tests of personality and intelligence. This method was used in the D analyses.

Summary. The scripts of 23 freshmen were analyzed by Mr. Lucas, a professional graphologist (L), and by Professor Downey (D). The experimenter was well acquainted with the subjects to whom he had given during the previous year a large number of psychological tests. He attempted to validate the graphological analyses against psychological data.

The experimenter examined whole groups of sketches at one time, and, later, considered smaller subgroups. In both cases he obtained a moderate degree of success, ranging from 1.98 to 3.59 times the probable error. The subjects, however, were not usually able to identify themselves or their friends, though some were much more successful than others.

The graphologist selected those subjects who, he believed, would stand highest on certain personality traits. Professor Downey selected, instead, subjects whom she expected to obtain the highest or lowest scores on a number of standard tests which the subjects had actually taken. Very irregular agreement was obtained, both the L and D selections being high on some traits and low on others. The accuracy of the D selections was statistically fairly significant, but that of the L selections was scarcely equal to its P.E.

Each point in the L and D sketches which could be adequately verified by the experimenter was examined. Some subjects were thus found to have been far more successfully analyzed than others, while certain qualities (physical characteristics, abilities, interests, social traits, etc.), showed a similar diversity of success. The average accuracy of all the points was +9.3% for the L and +15.7% for the D analyses.

The immaturity of the subjects, and the possible bias or lack of knowledge of the subjects on the part of the experimenter, together with certain other deficiencies of technique, may be in part responsible for the somewhat low (though uniform) degree of validation found by four out of the five methods employed.

Discussion. In conducting these experiments on handwriting and personality and in reporting them in Chapters X and XI, the authors have endeavored to maintain an impartial and scientific point of

view. They were aware of both the pointed and pointless criticisms made by psychologists against graphology, as well as of the exaggerations of some of its enthusiastic practitioners; but to the best of their abilities they resisted both types of prejudice.

Some of the conditions observed in these experiments which might have placed the graphologists at an unfair disadvantage have already been mentioned. In order to take into account the remaining criticisms and final opinions of the coöperating analysts the following brief discussion is added.

(1) Mr. Lucas writes concerning the results, "I am not disappointed, and I have not failed. If you are able to keep track of these subjects, you will find that with the passing of the years, my estimate of these immature young men will be proved far more accurate than your present perspective."

It is, of course, impossible to deny the graphologist's assertion that time alone will provide an adequate validation of the analyses. It is quite conceivable that germinal characteristics in personality are accessible to the graphologist and not to the psychologist or to the friendly acquaintance. This confusion of prophecy with contemporary description is often found in graphological analyses, and constitutes a serious obstacle in the way of immediate and objective validation.

(2) Besides stressing the immaturity of the subjects, Mr. Lucas calls attention to their homogeneity (all being freshmen in an eastern college) and to the absence of women in the group. Undoubtedly the lack of heterogeneity makes for failures in all matching experiments. The studies reported in both Chapter X and Chapter XI suffer somewhat under this handicap.

(3) It should also be pointed out that Mr. Lucas wrote sketches of considerable length. The subjective difficulties involved in matching or identifying his sketches were thereby greatly increased. It is therefore quite possible that this factor had an adverse influence upon the results. The authors incline to support Mr. Lucas in his suggestion that in the future when methods of matching or identification are employed the brief "cameo" sketch be used in place of the longer analyses. The optimum length and proper composition of brief case studies of personality for research involving the methods of matching and identification have not yet been determined.

(4) Mr. Lucas reminds us that half an hour or less is devoted to a graphological analysis, whereas psychological studies consume much

more time. Although this consideration has no bearing upon the final accuracy of the two methods, it does call attention to their relative economy. One might frame the following problem for future investigation: If only half an hour were devoted to the study of a personality, would the psychological or the graphological method have greater validity?

(5) Dr. Downey raises the interesting question whether pooling the results of analyses by graphologists and by psychologists might attain a higher validity than either method employed alone. This query is especially pertinent in the face of our inability to demonstrate that either method provides a criterion superior to the other. But this very lack of a suitable criterion makes all such problems of determining the diagnostic value of different methods, and combining them into a single battery, very complex. This is an issue for the future rather than for the present.

(6) The reader will have noticed in these experiments that the graphological analysts have submitted to rigid controls, that they have been forced to a large extent to abandon their natural way of working or else to submit their sketches to a strict and somewhat arbitrary scoring. It seems only fair, therefore, to conclude this discussion with a brief account of an additional "experiment" in which there were insufficient controls, but in which the graphologist was allowed every advantage.

Mr. Lucas made a long and searching analysis of the handwriting of one of the authors (without knowing his profession, name, or any facts whatsoever concerning him). The analysis, prepared by the graphologist in his own manner, mentioned 58 separate features of the writer's personality. Each feature was scored as "true," "partly true," or "false" by the writer himself and by two psychologists who were intimate acquaintances. Remarks which were ambiguous were considered as "true" if in any sense at all they could be applied to the writer. Using this lenient method of scoring only three of the 58 features were considered to be distinctly false by one or more of the judges. In other words, by this method, Mr. Lucas obtained a score of approximately 94% success. As unsatisfactory as this method is from the scientific point of view, it must be considered entirely typical of the conditions under which the graphologist naturally works. A detailed account of this study is given by Powers (*133*).

CHAPTER XII

SUMMARY AND CONCLUSIONS

1. The general interest in graphology in Europe and the serious experimentation there stand in marked contrast to the misapprehensions and neglect of the field among American psychologists. Commercial abuse and lack of scientific standards on the part of many professional graphologists are in part responsible for the state of affairs in this country, but lack of information and prejudice among psychologists are likewise significant factors.

2. On *a priori* grounds handwriting, which may be considered as "crystallized gesture," should be expected to furnish a valuable though intricate index to the patterns of personality. That deep-lying central dispositions in personality play a part in the determination of handwriting is indicated by several facts, such as the development of marked graphic individuality with age, the similarity of scripts written by the same individual using different groups of muscles, and the incidence of hereditary resemblances in handwriting.

3. The factors which determine the final form of the script, however, are exceedingly complex, and include the system of penmanship taught, the objective conditions of writing, the age of the writer, the spontaneity of his writing, his temporary mood and state of health, as well as his enduring personal make-up. Difficulties must, therefore, be expected in establishing correspondences between the script and the writer's personal traits.

4. Contemporary graphologists deal more frequently

with general features of script and their interrelations than with detailed elements or "signs." No adequate test of graphology, therefore, is afforded by experiments which are based on the measurement of specific and detailed "signs."

5. Though most graphologists rely to a greater or lesser extent on general "intuitive" impressions of the total form-quality of the script, the most successful systems seem to make some use of the analysis of actual movements of the hand, of variations in speed and pressure, and other information drawn from the direct study of the motor processes. Hence research in the psychology of handwriting by legal, pedagogical, and psychiatric investigators, as well as by psychologists themselves, is necessary for the development of scientific graphology.

6. In attempting to validate graphological analyses of personality, the problem of a suitable criterion offers much difficulty. Objective psychological tests and ratings are fragmentary; they tend to yield a mosaic of disconnected scores, while graphology claims to diagnose the underlying unities of the total personality. On the other hand the terminology employed by graphologists is loose and quite unsatisfactory to the psychologist. It is difficult to find a common meeting ground for an adequate and controlled testing of graphological skill.

7. Promising results have been obtained by Binet, Saudek, and others, through the *sorting* method. The scripts of two distinct groups of writers are shuffled together, and are then sorted into appropriate groups by the graphologists who are free to employ any method of analysis they choose. In this way Saudek claims to have sorted the handwritings of honest and dishonest people with nearly 100% success. Similarly in an experiment reported in Chapter XI, Professor Downey selected from

a group of 23 scripts, the writers who in her estimation would obtain distinctly high or low scores on certain standard psychological tests. Though her success varied considerably with respect to different tests, yet her average results were to a fairly significant degree superior to chance.

8. Many investigators have applied this method of sorting to the detection of sex from handwriting. These experiments are sometimes cited as unfavorable to graphology, but this conclusion demonstrates the unfamiliarity of many psychologists with graphological claims. Scarcely any graphologist maintains that sex can be detected correctly in more than 60% to 70% of the cases.

9. An important validatory technique is the *matching* method wherein a small group of anonymous scripts, or of graphological analyses based on these scripts, is compared with a set of psychological case studies. A number of judges attempt to match each specimen or analysis with the appropriate case study, or, when the subjects are known to them, to identify the graphological sketches directly. In one study by Bobertag, 80% success was achieved by untrained judges, where only 20% could have been obtained by chance.

10. There are certain disadvantages in the method of matching. Chief among these are its dependence upon the wide differences in the ability of individual judges to match, upon the accuracy of their knowledge of the subjects, and upon certain accidents in respect to the order in which the matchings are attempted. Furthermore, since this method interposes the judgment of a third party, it is not a direct test of the graphologist's skill. With certain refinements, however, the method is a useful one, lending itself to virtually every scientific control which the psychologist desires to employ.

11. Two series of matching experiments are described in detail in Chapters X and XI. In both the success of the judges was superior to chance, though by only a small amount. It is shown, however, that conditions were not entirely suitable to a completely adequate validation, and that the results should not therefore be regarded as a final measure of the dependability of graphological methods.

12. Although the results are not greatly above chance, yet their uniformity (whatever the method used), together with the prevailingly "good errors" and occasional brilliant successes, suggest even higher validity which, owing to methodological inadequacies, remains so far unestablished.

13. The results show that practicing graphologists seem on the whole to give better judgments and to make more correct matchings than do untrained persons. Even at best, however, the results do not establish professional graphology on a certain basis; by our criteria the results are still nearer to chance than to perfection. But in favor of the graphologists it may be urged that psychological criteria are far from perfect, that many items in an analysis may be predictions which time alone can affirm or deny, that the conditions of our experiments have been so stringently controlled as to be prejudicial to the free exercise of the graphologists' talents.

14. In seeking the coöperation of graphologists in these experiments the writers have found in general a willingness and open-mindedness that stand in contrast to the prevailing view that all experts in handwriting are guilty of meretriciousness and deceit.

15. Regarding the value of commercial graphology the authors have no final opinion. Probably too much has been claimed for it by graphologists and too little by

psychologists. Practically all lines of research reported in the last three chapters (excepting those that employ detailed "signs") have been favorable to a slight degree. The success of professional graphologists seems usually to exceed chance, and occasionally to show brilliance. Some experts appear to give better readings than others, although each seems to be irregular and variable in the aptness of his judgments. It is certain that there is much room for improvement in graphological practice, both through the enhancement of individual skill and through advancing the ethics of the practice.

16. In Part A of this volume the *elements* of handwriting (size, speed, point, and grip pressures) were found to correlate with many attributes of movement selected from widely different performances. In Part B, the *pattern* of handwriting, its total graphic character, was found likewise to be interlocked with other expressive behavior. Judgments made from script and judgments made from the direct observation of behavior showed a definite, though not perfect, correspondence.

17. The relation of these studies in handwriting to the argument of this volume should now be apparent. Neither handwriting nor any other phenomenon of movement that we have studied is entirely unrelated to other motor activity. Although the evidence is often complex, and by no means complete, our work on handwriting points to intricate psychosomatic organization, just as did our work on other aspects of gesture.

In general we feel that our methods with their rigid controls have served to establish the minimum, rather than the maximum, consistency between the activities that we have studied. The evidence indicates clearly that the expressive movements of personality are not specific and unrelated; on the contrary they form coherent, if

perplexing, patterns. Many of these patterns have been disclosed by our methods; others undoubtedly remain to be traced.

From our results it appears that a man's gesture and handwriting both reflect an essentially stable and constant individual style. His expressive activities seem not to be dissociated and unrelated to one another, but rather to be organized and well-patterned. Furthermore, the evidence indicates that there is congruence between expressive movement and the attitudes, traits, values, and other dispositions of the "inner" personality. Although the question of the organization of the personality as a whole is beyond the scope of this volume, it is clear that the foundations for an adequate solution of this important problem cannot be supplied by the anarchic doctrine of specificity but only by the positive and constructive theories of consistency.

REFERENCES

1. Allen, C. N., Individual differences in delayed reaction of infants. *Arch. of Psychol.*, 1931, Vol. 19, 5–40.
2. Allport, F. H., *Social Psychology.* Boston: Houghton Mifflin, 1924.
3. Allport, G. W., *An Experimental Study of the Traits of Personality* (unpublished). Harvard College Library, 1922.
4. Allport, G. W., and Vernon, P. E., The field of personality. *Psychol. Bull.*, 1930, Vol. 27, 677–730.
5. Anders, P., Ueber die individuellen Eigenrhythmus beim menschlichen Gange und seine Beziehungen zum Rhythmus der Herz- und Atemtätigkeit. *Pflüg. Arch.*, 1928, Vol. 220, 287–299.
6. Anderson, V. V., *Psychiatry in Industry.* New York: Harper and Bros., 1929.
7. Arnheim, R., Experimentell-psychologische Untersuchungen zum Ausdrucksproblem. *Psychol. Forsch.*, 1928, Vol. 11, 1–132.
8. Baxter, M. F., An experimental study of the differentiation of temperament on the basis of rate and strength. *Amer. J. Psychol.*, 1927, Vol. 38, 59–96.
9. Bernstein, E., Quickness and intelligence. *Brit. J. Psychol. Monog. Suppl.*, 1924, Vol. 3, No. 7.
10. Bläsch, H., Zur Graphologie. *Psychol. Rundschau,* 1929, Vol. 1, 201–204.
11. Bläsch, H., Graphologische Bemerkungen über die Bindungsform. *Psychol. Rundschau,* 1929, Vol. 1, 230–234.
12. Bills, A. G., The influence of muscular tension on the efficiency of mental work. *Amer. J. Psychol.*, 1927, Vol. 38, 227–251.
13. Bills, A. G., Tension in learning and association. *Proc. IX Internat. Cong. Psychol.*, pp. 75–76. Princeton: Psychol. Rev. Co., 1930.
14. Bills, A. G., and Brown, C., The quantitative set. *J. Exper. Psychol.*, 1929, Vol. 12, 301–323.
15. Binet, A., *Les Révélations de l'Écriture d'après un Controle Scientifique.* Paris: Alcan, 1906.
16. Bobertag, O., *Ist die Graphologie Zuverlässig?* Heidelberg: Kampmann, 1929.

17. Bobertag, O., Bermerkungen zur Praxis und Theorie der Graphologien. *Zsch. f. Menschenkenntnis*, 1930, Vol. 6, 300–317.
18. Bogen, H., and Lippmann, O., Gang und Charakter. *Beihefte zur Zsch. f. ang. Psychol.*, No. 58. Leipzig: Barth, 1931.
19. Boven, W., *La Science du Caractère.* Neuchatel: Delachaux et Niestle, S. A., 1931.
20. Braun, F., Untersuchungen über das persönliche Tempo. *Arch. f. d. ges. Psychol.*, 1927, Vol. 60, 317–360.
21. Bray, C. W., Transfer of learning. *J. Exper. Psychol.*, 1928, Vol. 11, 443–467.
22. Brooks, C. H., *Your Character from Your Handwriting.* London: Allen and Unwin, 1930.
23. Broom, M. E., Thompson, B., and Bouton, M. T., Sex differences in handwriting. *J. Appl. Psychol.*, 1929, Vol. 13, 159–166.
24. Bühler, C., The social behavior of the child, in *Handbook of Child Psychology*, pp. 392–431. Worcester: Clark Univ. Press, 1931.
25. Bührig, W., Einiges über den Schriftdruck. *Psychol. Rundschau*, 1930, Vol. 2, 270–276.
26. Burt, C. L., *The Measurement of Mental Capacities.* Edinburgh: Oliver and Boyd, 1927.
27. Campbell, M., Individual differences in the speed of serial reaction of large-muscle groups. *Psychol. Bull.*, 1930, Vol. 27, 639.
28. Cleeton, G. U., and Knight, F. B., Validity of character judgments based on external criteria. *J. Appl. Psychol.*, 1924, Vol. 8, 215–231.
29. Coghill, G. E., *Anatomy and the Problem of Behavior.* New York: Macmillan, 1929.
30. Cooley, C. H., *Human Nature and the Social Order.* New York: C. Scribner's Sons, 1902.
31. Couvé, R., Graphologische Berufseignungsuntersuchungen. *Industr. Psychotech.*, 1926, Vol. 3, 114–117.
32. Crépieux-Jamin, J., *The Psychology of the Movements of Handwriting.* Trans. and arranged by L. K. Given-Wilson. London: Routledge, 1926.
33. Delabarre, E. B., Logan, R. R., and Reed, A. Z., The force and rapidity of reaction movement. *Psychol. Rev.*, 1897, Vol. 4, 615–631.

34. Delhougne, A., Dreidimensionale Graphologie. *Zsch. f. Menschenkenntnis*, 1931, Vol. 7; *Zentrallblatt f. Graph.*, 1–16.
35. Diehl, A., Ueber die Eigenschaften der Schrift bei Gesunden. *Psychol. Arbeiten*, 1899, Vol. 3, 1–61.
36. Dodge, R., *Conditions and Consequences of Human Variability*. New Haven: Yale Univ. Press, 1931.
37. Downey, J. E., Preliminary study of family resemblances in handwriting. *Univ. Wyom. Dept. Psychol. Bull.*, No. 1. Laramie: Univ. of Wyoming Press, 1910.
38. Downey, J. E., Judgments on the sex of handwriting. *Psychol. Rev.*, 1910, Vol. 17, 205–216.
39. Downey, J. E., *Graphology and the Psychology of Handwriting*. Baltimore: Warwick and York, 1919.
40. Downey, J. E., *The Will-Temperament and Its Testing*. Yonkers, N. Y.: World Book Co., 1923.
41. Drever, J., Notes on the experimental study of writing. *J. Exper. Ped.*, 1913, Vol. 2, 25–29.
42. Drever, J., The analytical study of the mechanism of writing. *Proc. Royal Soc. Edinburgh*, 1913–14, Vol. 34, 230–240.
43. Duffy, E., Tensions and emotional factors in reaction. *Genet. Psychol. Monog.*, 1930, Vol. 7, No. 1.
44. Duffy, E., The measurement of muscular tension as a technique for the study of emotional tendencies. *Amer. J. Psychol.*, 1932, Vol. 44, 146–162.
45. Dunlap, K., *The Elements of Scientific Psychology*. St. Louis: C. V. Mosby, 1922.
46. Dunlap, K., The rôle of eye-muscles and mouth-muscles in the expression of emotion. *Genet. Psychol. Monog.*, 1927, Vol. 2, No. 3.
47. Durkheim, E., Répresentations individuelles et répresentations collectives. *Rev. de métaphys.*, 1898, Vol. 6, 274–302.
48. Enke, W., Die Psychomotorik der Konstitutionstypen. *Zsch. f. ang. Psychol.*, 1930, Vol. 36, 237–287.
49. Flach, A., Die Psychologie der Ausdrucksbewegung. *Arch. f. d. ges. Psychol.*, 1928, Vol. 65, 435–534.
50. Von Foerster, J. F., Verfahren und Möglichkeiten der Schriftbeurteilung. *Industr. Psychotech.*, 1927, Vol. 4, 129–147.
51. Ford, A., Attention automatization. *Amer. J. Psychol.*, 1929, Vol. 41, 1–32.
52. Freeman, F. N., An experimental analysis of the writing movement. *Psychol. Monog.*, 1914, Vol. 17, No. 4.

53. FREEMAN, G. L., The measurement of tonus by deformation of the tendon. *Amer. J. Psychol.*, 1930, Vol. 42, 581–592.
54. FREEMAN, G. L., The spread of neuro-muscular activity during mental work. *J. Gen. Psychol.*, 1931, Vol. 5, 479–492.
55. FREUD, S., *The Psychopathology of Everyday Life.* New York: Macmillan, 1914.
56. GALTON, F., *Inquiries into Human Faculty and Its Development.* London: Macmillan, 1883.
57. GARRETT, H. E., Personality as "habit organization." *J. Abn. and Soc. Psychol.*, 1926, Vol. 21, 250–255.
58. GESELL, A. L., Accuracy in handwriting as related to school intelligence and sex. *Amer. J. Psychol.*, 1906, Vol. 17, 394–405.
59. GESELL, A. L., The developmental psychology of twins, in *Handbook of Child Psychology*, pp. 158–203. Worcester: Clark Univ. Press, 1931.
60. GOLLA, F. L., and ANTINOVITCH, S., The relation of muscle tonus and the patellar reflex to mental work. *J. Ment. Sci.*, 1929, Vol. 75, 234–241.
61. GROSS, A., Untersuchungen über die Schrift Gesunder und Geisteskranker. *Psychol. Arbeiten*, 1899, Vol. 2, 450–567.
62. GROSS, H., *Criminal Psychology.* Transl. from 4th German edit. Boston: Little, Brown, 1918.
63. GUREVICH, M., Motorik, Körperbau, und Charakter. *Arch. f. Psychiatr. u. Nervenkr.*, 1926, Vol. 76, 521–532.
64. GUTTMANN, A., Das Tempo und seine Variationsbreite. *Zsch. f. ang. Psychol.*, 1931, Vol. 40, 65.
65. HALVERSON, H. M., An experimental study of prehension in infants by means of systematic cinema records. *Genet. Psychol. Monog.*, 1931, Vol. 10, Nos. 2 and 3.
66. HARTSHORNE, H., MAY, M. A., and SHUTTLEWORTH, F. K., *Studies in the Organization of Character.* New York: Macmillan, 1930.
67. HENNING, H., Ziele und Möglichkeiten der experimentellen Charakterprüfung. *Jahrb. d. Charakterol.*, 1929, Vol. 6, 213–274.
68. HIRT, E., Untersuchungen über das Schreiben und die Schrift. *Psychol. Arbeiten*, 1914, Vol. 6, 531–664.
69. HOLLINGWORTH, H. L., Correlations of achievement within an individual. *J. Exper. Psychol.*, 1925, Vol. 8, 190–208.

70. HOLLINGWORTH, H. L., *Vocational Psychology and Character Analysis.* New York: Appleton, 1929.
71. HOLT, E. B., *Animal Drive and the Learning Process.* New York: Henry Holt, 1931.
72. HSIAO, H. H., Is Thorndikianism the last word in the psychology of learning? *J. Gen. Psychol.*, 1929, Vol. 2, 115–129.
73. HÜBEL, W., Ueber psychische Geschwindigkeiten und ihre gegenseitigen Beziehungen. *Zsch. f. ang. Psychol.*, 1930, Vol. 35, 447–496.
74. HULL, C. L., *Aptitude Testing.* Yonkers, N. Y.: World Book Co., 1928.
75. HULL, C. L., and MONTGOMERY, R. P., Experimental investigation of certain alleged relations between character and handwriting. *Psychol. Rev.*, 1919, Vol. 26, 63–74.
76. ICHHEISER, G., Die Ueberschätzung der Einheit der Persönlichkeit als Täuschungsquelle bei der psychologischen Beurteilung der Menschen. *Zsch. f. ang. Psychol.*, 1929, Vol. 33, 273–287.
77. ICHHEISER, G., Das Du-bewusstsein und seine ethisch-ästhetische Struktur. *Zsch. f. ang. Psychol.*, 1930, Vol. 36, 369–392.
78. JACOBSON, E., *Progressive Relaxation.* Chicago: Univ. of Chicago Press, 1929.
79. JOHNSON, B., Changes in muscular tension in coördinated hand movements. *J. Exper. Psychol.*, 1928, Vol. 11, 329–341.
80. JOHNSON, H. M., and SWAN, T. H., Sleep. *Psychol. Bull.*, 1930, Vol. 27, 1–39.
81. KEHR, T., Versuchsanordnung zur experimentellen Untersuchung einer kontinuerlichen Aufmerksamkeitsleistung. *Zsch. f. ang. Psychol.*, 1916, Vol. 11, 465–480.
82. KELLEY, T. L., *Statistical Method.* New York: Macmillan, 1923.
83. KEMPF, E. J., *The Autonomic Functions and the Personality.* New York: Nervous and Mental Disease Publ. Co., 1921.
84. KINDER, J. S., A new investigation of judgments on the sex of handwriting. *J. Educ. Psychol.*, 1926, Vol. 17, 341–344.
85. KLAGES, L., *Ausdrucksbewegung und Gestaltungskraft*, 3d edit. Leipzig: Barth, 1923.
86. KLAGES, L., *Handschrift und Charakter*, 13th edit. Leipzig: Barth, 1929.
87. KLAGES, L., *The Science of Character.* Transl. from 5th edit. *Grundlagen der Charakterkunde.* London: Allen and Unwin, 1929.

88. Köhler, W., Nachweis einfacher Strukturfunktionen beim Schimpansen und beim Haushuhn. *Abhandl. d. Preuss. Akad. d. Wissenschaft*, 1918, *Phys.-Math.*, Kl. Nr. 2.
89. Köhler, W., *Gestalt Psychology*. New York: Liveright, 1929.
90. Köhler, W., Les forces motrices du comportement. *Jl. de psychol.*, 1930, Vol. 27, 367–390.
91. Kramer, E., and Lauterbach, C. E., Resemblance in the handwriting of twins and siblings. *J. Educ. Res.*, 1928, Vol. 18, 149–152.
92. Krauss, R., Ueber graphischen Ausdruck. *Beihefte zur Zsch. f. ang. Psychol.*, No. 48. Leipzig: Barth, 1930.
93. Krout, M. H., A preliminary note on some obscure symbolic muscular responses of diagnostic value in the study of normal subjects. *Amer. J. Psychiatr.*, 1931, Vol. 11, 29–71.
94. Von Kügelgen, G., Graphologie und Berufseignung. *Industr. Psychotech.*, 1928, Vol. 5, 311.
95. Land, A. H., Graphology, a psychological analysis. *Univ. of Buffalo Stud.*, 1924, Vol. 3, 81–114.
96. Landis, C., Gullette, R., and Jacobsen, C., Criteria of emotionality. *Ped. Sem.*, 1925, Vol. 32, 209–234.
97. Lankes, W., Perseveration. *Brit. J. Psychol.*, 1914–15, Vol. 7, 387–419.
98. Lashley, K. S., *Brain Mechanisms and Intelligence*. Chicago: Univ. of Chicago Press, 1929.
99. Lersch, P., Die Bedeutung der mimischen Ausdruckserscheinungen für die Beurteilung der Persönlichkeit. *Industr. Psychotech.*, 1928, Vol. 5, 178–183.
100. Lewin, K., Kindlicher Ausdruck. *Zsch. f. päd. Psychol.*, 1927, Vol. 28, 510–526.
101. Lewin, K., Environmental forces in child development and behavior, in *Handbook of Child Psychology*, pp. 94–127. Worcester: Clark Univ. Press, 1931.
102. Lewis, N. D. C., The practical value of graphic art in personality studies. *Psychoan. Rev.*, 1925, Vol. 12, 316–322.
103. Lewitan, C., Untersuchungen über das allgemeine psychomotorische Tempo. *Zsch. f. Psychol.*, 1927, Vol. 101, 321–376.
104. Mayer-Benz, L., Schriftdeutung jugendlicher Psychopathen. *Zsch. f. Kinderforsch.*, 1929, Vol. 35, 637–649.
105. Melcher, W. A., Dual personality in handwriting. *J. Crim. Law and Criminol.*, 1920, Vol. 11, 209–216.

106. MELOUN, J., Experimentelle Graphologie. *Zsch. f. ang. Psychol.*, 1929, Vol. 32, 518–525.
107. MELOUN, J., Objektive Kontrollmethoden in der Schriftpsychologie. *Arch. f. d. ges. Psychol.*, 1929, Vol. 71, 357–368.
108. MEUMANN, E., *Vorlesungen zur Einführung in die experimentelle Pädagogik*, 2d edit. Leipzig: Engelmann, 1913.
109. MEYER, G., *Die wissenschaftlichen Grundlagen des Schreibens*, 2d edit. Jena, 1925.
110. MILES, W. R., Correlation of reaction and coördination speed with age in adults. *Amer. J. Psychol.*, 1931, Vol. 43, 377–391.
111. MORF, G., Ein graphologischer Leckerbissen. *Psychol. Rundschau*, 1930, Vol. 2, 176–177.
112. MORGAN, C. L., *Emergent Evolution.* London: Williams and Norgate, 1923.
113. MORGAN, J. J. B., The overcoming of distraction and other resistances. *Arch. of Psychol.*, 1916, Vol. 5, No. 35.
114. MORGAN, J. J. B., *The Psychology of Abnormal People.* New York: Longmans, Green, 1928.
115. MÜHL, A. M., Automatic writing as an indication of the fundamental factors underlying the personality. *J. Abn. and Soc. Psychol.*, 1922–23, Vol. 17, 162–183.
116. MÜNSTERBERG, H., Die psychophysische Grundlage der Gefühle. *Proc. II Internat. Cong. of Exper. Psychol.*, pp. 132–135. London: Williams and Norgate, 1892.
117. MUNN, N. L., Bilateral transfer of learning. *J. Exper. Psychol.*, 1932, Vol. 15, 343–353.
118. MURPHY, G., and L. B., *Experimental Social Psychology.* New York: Harper and Bros., 1931.
119. MYERS, C. S., Psychological cautions in the use of statistics. *Zsch. f. ang. Psychol.*, 1930, Vol. 36, 82–86.
120. MCALLISTER, C. N., Researches on movements used in handwriting. *Yale Psychol. Lab. Stud.*, 1900, Vol. 8, 21–63.
121. MCKINLEY, J. C., and BERKWITZ, N. J., Quantitative studies on human muscle tonus: I, Description of methods. *Arch. Neurol. and Psychiatr.*, 1928, Vol. 19, 1036–1056.
122. NEWHALL, S. M., Sex differences in handwriting. *J. Appl. Psychol.*, 1926, Vol. 10, 151–161.
123. NORCROSS, W. H., Experiments on the transfer of training. *J. Comp. Psychol.*, 1921, Vol. 1, 317–363.
124. OLSON, W. C., The measurement of nervous habits in normal

children. *Instit. Child Wel. Monog. Ser.*, No. 3. Minneapolis: Univ. of Minnesota Press, 1929.

125. Omwake, K. T., The value of photographs and handwriting in estimating intelligence. *Publ. Personnel Stud.*, 1925, Vol. 3, 2–15.
126. Osborn, A. S., *Questioned Documents*, 2d edit. Albany: Boyd Printing Co., 1929.
127. Oseretzky, N., Psychomotorik. *Beihefte zur Zsch. f. ang. Psychol.*, No. 57. Leipzig: Barth, 1931.
128. Paterson, D. G., *Physique and Intellect.* New York: Century Co., 1930.
129. Paterson, D. G., and Ludgate, K. E., Blonde and brunette traits: a quantitative study. *J. Personnel Res.*, 1922, Vol. 1, 122–128.
130. Pear, T. H., *Voice and Personality.* London: Chapman and Hall, 1931.
131. Perrin, F. A. C., Physical attractiveness and repulsiveness. *J. Exper. Psychol.*, 1921, Vol. 4, 203–217.
132. Ponder, E., and Kennedy, W. P., On the act of blinking. *Quart. J. Exper. Physiol.*, 1927, Vol. 18, 89–110.
133. Powers, E., *Graphic Factors in Relation to Personality: an Experimental Study* (unpublished). Dartmouth College Library, 1930.
134. Pratt, K. C., Nelson, A. K., and Sun, K. H., *The Behavior of the Newborn Infant.* Columbus: Ohio Univ. Press, 1930.
135. Preyer, W., *Zur Psychologie des Schreibens*, 3d edit. Leipzig: L. Voss, 1928.
136. Prinzhorn, H., Bildnerei der Geisteskranken. Berlin: Springer, 1922.
137. *Proceedings of the Second Colloquium on Personality Investigation.* Baltimore: Johns Hopkins Univ. Press, 1930.
138. Ream, M. J., The tapping test: a measure of motility. *Psychol. Monog.*, 1922, Vol. 31, No. 1, pp. 293–319.
139. Reymert, M. L., The personal equation in motor capacities. *Scandin. Sci. Rev.*, 1923, Vol. 2, 177–222.
140. Rice, L., *Character Reading from Handwriting.* New York: Stokes, 1927.
141. Roback, A. A., Meaning in personality manifestations. *Proc. IX Internat. Cong. of Psychol.*, pp. 362 f. Princeton: Psychol. Rev. Co., 1930.

142. Roback, A. A., *Personality: the Crux of Social Intercourse.* Cambridge, Mass.: Sci-Art, 1931.
143. Román-Goldzieher, K., Untersuchungen über die Schrift der Stotterer, Stammler, und Polterer. *Zsch. f. Kinderforsch.*, 1929, Vol. 35, 116–139.
144. Román-Goldzieher, K., "Graphodyn." *Zsch. f. ang. Psychol.*, 1931, Vol. 40, 23–33.
145. Ruch, G. M., Ackerson, L., and Jackson, J. D., An experimental study of the Spearman-Brown formula as applied to educational test material. *J. Educ. Psychol.*, 1926, Vol. 17, 309–313.
146. Ruttman, W. J., Ueber Ausdrucksprüfung. *Zsch. f. päd. Psychol.*, 1915, Vol. 16, 131–136.
147. Sapir, E., Speech as a personality trait. *Amer. J. Sociol.*, 1927, Vol. 32, 892–905.
148. Saudek, R., *The Psychology of Handwriting.* London: Allen and Unwin, 1925.
149. Saudek, R., The methods of graphology. *Brit. J. Med. Psychol.*, 1927, Vol. 7, 221–259.
150. Saudek, R., *Experiments with Handwriting.* New York: Morrow, 1928.
151. Saudek, R., Das zentrale Nervensystem und der Schreibakt. *Jahrb. d. Charakterologie*, 1929, Vol. 6, 277–303.
152. Saudek, R., Zur psychodiagnostischen Ausdeutung des Schreibdrucks. *Zsch. f. ang. Psychol.*, 1931, Vol. 39, 433–449.
153. Saudek, R., Zur Psychologie der amerikanischen Handschrift. *Zsch. f. Menschenkenntnis*, 1931, Vol. 6; *Zentrallblatt f. Graph.*, 361–396.
154. Saudek, R., Writing movements as indications of the writer's social behavior. *J. Soc. Psychol.*, 1931, Vol. 2, 337–373.
155. Saudek, R., Pubertätsjahre im englischen Internat. *Charakter*, 1932, Vol. 1, 12–26.
156. Saunders, E. B., and Isaacs, S., Tests of reaction time and motor inhibition in the psychoses. *Amer. J. Psychiatr.*, 1929–30, Vol. 9, 79–112.
157. Schaeffer, A. A., Spiral movement in man. *J. Morph.*, 1928, Vol. 45, 293–398.
158. Scheck, M. G., Involuntary tongue movements under varying stimuli. *Proc. Iowa Acad. Sci.*, 1925, Vol. 32, 385–391.
159. Schorn, M., Untersuchungen zur Kritik der graphologischen Gutachten. *Industr. Psychotech.*, 1927, Vol. 4, 359–368.

160. SCHRIJVER, J., Een Graphologisch-controle-onderdoek. *Tydschrift voor Wetenschappelyke Graphologie*, 1929, Vol. 1, 117–128.
161. SCHWANGART, F., Persönliches bei Tieren und tierisches Niveau. *Jahrb. d. Charakterologie*, 1928, Vol. 5, 103–140.
162. SCRIPTURE, E. W., *The Elements of Experimental Phonetics.* New York: Scribner's Sons, 1902.
163. SEELIG, E., Die Registrierung unwillkürlicher Ausdrucksbewegungen als forensischpsychodiagnostische Methode. *Zsch. f. ang. Psychol.*, 1927, Vol. 28, 45–84.
164. SEESEMANN, K., Bewährungskontrolle graphologischer Gutachten. *Industr. Psychotech.*, 1929, Vol. 6, 104–108.
165. SHERMAN, I. C., The Franz dot tapping test as a measure of attention. *J. Appl. Psychol.*, 1923, Vol. 7, 353–359.
166. SMITH, M., CULPIN, M., and FARMER, E., A study of telegraphists' cramp. *Industr. Fat. Res. Board Rep.*, No. 43. London: H. M. Stat. Office, 1927.
167. SPEARMAN, C. E., Correlations of sums or differences. *Brit. J. Psychol.*, 1912–13, Vol. 5, 417–426.
168. SPEARMAN, C. E., *The Abilities of Man.* London: Macmillan, 1927.
169. SPRANGER, E., *Types of Men.* Transl. from 5th German edit. of *Lebensformen.* Halle: Niemeyer, 1928.
170. STARCH, D., The measurement of efficiency in writing. *J. Educ. Psychol.*, 1915, Vol. 6, 106–114.
171. STARCH, D., The similarity of brothers and sisters in mental traits. *Psychol. Rev.*, 1917, Vol. 24, 235–238.
172. STERN, W., *Person und Sache: I, Ableitung und Grundlehre*, 3d edit. Leipzig: Barth, 1923.
173. STERN, W., *Person und Sache: III, Wertphilosophie.* Leipzig: Barth, 1924.
174. STERN, W., Review of works by L. Klages. *Zsch. f. ang. Psychol.*, 1926, Vol. 27, 268–275.
175. STERN, W., *Studien zur Personwissenschaft: I, Personalistik als Wissenschaft.* Leipzig: Barth, 1930.
176. STRATTON, G. M., *Developing Mental Power.* Boston: Houghton Mifflin, 1922.
177. STROUD, J. B., Apparatus for measuring muscular tension. *J. Exper. Psychol.*, 1931, Vol. 14, 184–185.
178. SYMONDS, P. M., *Diagnosing Personality and Conduct.* New York: Century Co., 1932.

179. THEISS, H., Experimentelle Untersuchungen über die Erfassung des handschriftlichen Ausdrucks durch Laien. *Psychol. Forsch.*, 1931, Vol. 15, 276–358.
180. THOMAS, W. I., and D. S., *The Child in America.* New York: Knopf, 1928.
181. THORNDIKE, E. L., Handwriting. *Teach. Coll. Rec.*, 1910, Vol. 11, 83–175.
182. THORNDIKE, E. L., *The Principles of Teaching.* New York: A. G. Seiler, 1906.
183. THURSTONE, L. L., *The Reliability and Validity of Tests.* Ann Arbor: Edward Bros., 1931.
184. UHRBROCK, R. S., An analysis of the Downey Will-Temperament Tests. *Teach. Coll. Contrib. Educ.*, No. 296. New York: Columbia Univ. Press, 1928.
185. VANCE, J. G., *A Mirror of Personality.* London: Williams and Norgate, 1927.
186. VERNON, P. E., The personality of the composer. *Music and Letters*, 1930, Vol. 11, 58–70.
187. VERNON, P. E., Some characteristics of the good judge of personality. *J. Soc. Psychol.*, 1933, Vol. 4.
188. VOLLMER, A., Revision of the Atcherley *modus operandi* system. *J. Crim. Law and Criminol.*, 1919, Vol. 10, 229–274.
189. WAGONER, L. C., Speech as an indication of temperamental traits. *Quart. J. Speech Educ.*, 1925, Vol. 11, 237–242.
190. WAGONER, L. C., and ARMSTRONG, E. M., The motor control of children as involved in the dressing process. *J. Genet. Psychol.*, 1928, Vol. 35, 84–97.
191. WAGONER, L. C., and DOWNEY, J. E., Speech and Will-Temperament. *J. Appl. Psychol.*, 1922, Vol. 6, 291–297.
192. WASHBURN, R. W., A study of the smiling and laughing of infants in the first year of life. *Genet. Psychol. Monog.*, 1929, Vol. 6, Nos. 5–6.
193. WATSON, J. B., *Psychology from the Standpoint of a Behaviorist.* Philadelphia: Lippincott, 1919.
194. WELLS, F. L., Normal performance in the tapping test. *Amer. J. Psychol.*, 1908, Vol. 19, 437–483.
195. WELLS, F. L., Studies in retardation. *Amer. J. Psychol.*, 1909, Vol. 20, 38–59.
196. WILSON, E. B., Mathematics and statistics. *J. Amer. Stat. Assoc.*, 1930, Vol. 25, N. S. 169, 1–8.

197. WITTLICH, B., Ein Beitrag zur Frage der Geschlechtsbestimmung nach der Handschrift. *Zsch. f. Menschenkenntnis*, 1927, Vol. 3, 42–45.
198. WOLFE, H. K., Some judgments on the size of familiar objects. *Amer. J. Psychol.*, 1898, Vol. 9, 137–166.
199. WOLFF, W., Gestaltidentität in der Charakterologie. *Psychol. u. Med.*, 1929, Vol. 4, 32–44.
200. WOLFF, W., Ueber Faktoren charakterologischen Urteilsbildung. *Zsch. f. ang. Psychol.*, 1930, Vol. 35, 385–446.
201. WOODWORTH, R. S., *Dynamic Psychology*. New York: Columbia Univ. Press, 1918.
202. WOODWORTH, R. S., *Psychology, a Study of Mental Life*, rev. edit. New York: Henry Holt, 1929.
203. YIZLIN, S. G., Konstitution und Motorik: I, Zur Psychomotorik der Kretschmerschen Typen. *Zsch. f. d. ges. Neurol. u. Psychiatr.*, 1926, Vol. 105, 433–447.
204. YULE, G. U., *An Introduction to the Theory of Statistics*, 8th edit. London: Griffin, 1927.
205. ZIEHEN, T., *Die Grundlagen der Charakterologie*. Langensalza: Beyer, 1930.
206. ZOEPFFEL, H., Ein Versuch zur experimentellen Feststellung der Persönlichkeit im Säuglingsalter. *Zsch. f. Psychol.*, 1929, Vol. 111, 273–306.

INDEX OF NAMES

INDEX OF SUBJECTS